CANADIAN THESES ON MICROFICHE

THÈSES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

NL 339 (r. 88/09)

National Library of Canada / Bibliothèque nationale du Canada

Ottawa, Canada
K1A 0N4

61

TC –

ISBN 0-315-21262-4

CANADIAN THESES ON MICROFICHE SERVICE – SERVICE DES THÈSES CANADIENNES SUR MICROFICHE

PERMISION TO MICROFILM – AUTORISATION DE MICROFILMER

• Please print or type – Écrire en lettres moulées ou dactylographier

AUTHOR – AUTEUR

Full Name of Author – Nom complet de l'auteur

RONALD JACK HALL

Date of Birth – Date de naissance

MARCH 11, 1955

Canadian Citizen – Citoyen canadien

☑ Yes / Oui ☐ No / Non

Country of Birth – Lieu de naissance

CANADA

Permanent Address – Résidence fixe

12107 – 146 AVENUE
EDMONTON, ALBERTA
T5X-1V1

THESIS – THÈSE

Title of Thesis – Titre de la thèse

AN ANALYSIS OF SIMULATED THEMATIC MAPPER DIGITAL DATA FOR FOREST AND LAND COVER CLASSIFICATION IN THE NORTHWEST TERRITORIES.

Degree for which thesis was presented
Grade pour lequel cette thèse fut présentée

M.Sc.

Year this degree conferred
Année d'obtention de ce grade

1985

University – Université

UNIVERSITY OF ALBERTA

Name of Supervisor – Nom du directeur de thèse

DR. PETER CROWN

AUTHORIZATION – AUTORISATION

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to **microfilm** this thesis and to **lend** or **sell copies of the film.**

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de **microfilmer** cette thèse et de **prêter** ou de **vendre des exemplaires du film.**

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

ATTACH FORM TO THESIS – VEUILLEZ JOINDRE CE FORMULAIRE À LA THÈSE

Signature

Ronald J. Hall

Date

April 17, 1985.

NL-91 (r. 84/03)

Canada

THE UNIVERSITY OF ALBERTA

An Analysis of Simulated Thematic Mapper Digital Data for Forest and Land Cover Classification in the Northwest Territories

by

Ronald Jack Hall

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF Master of Science

IN

Remote Sensing

Departments of Soil Science and Forest Science

EDMONTON, ALBERTA

Spring, 1985

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Ronald Jack Hall

TITLE OF THESIS An Analysis of Simulated Thematic Mapper Digital Data for Forest and Land Cover Classification in the Northwest Territories

DEGREE FOR WHICH THESIS WAS PRESENTED Master of Science

YEAR THIS DEGREE GRANTED 1985

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Ronald J. Hall

PERMANENT ADDRESS:

12107 - 146 Avenue

Edmonton, Alberta

T5X-1V1

DATED April 17 1985.

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled An Analysis of Simulated Thematic Mapper Digital Data for Forest and Land Cover Classification in the Northwest Territories submitted by Ronald Jack Hall in partial fulfilment of the requirements for the degree of Master of Science in Remote Sensing.

.....

Supervisor

.....

Supervisor

W. A. Davis

.....

Date.....April 3, 1985.....

DEDICATION

"To Dad"

ABSTRACT

Methodology for broad forest classification was developed for simulated Thematic Mapper data. The study area in the Northwest Territories, offered unique challenges with respect to vegetation complexity, and number of prominent features on which to base the geometric correction of the digital data. Aerial photographs and radial line triangulation were employed to assist with the location of ground control points.

The unsupervised approach to image classification was considered to be the most feasible since there was little or no prior knowledge on class structure. The difficulties of determining the appropriate number of land cover classes and their content associated with this approach, were addressed by multivariate cluster analyses. This technique was assisted by a geographic referencing procedure based on an affine transformation, and SPSS on the geometrically corrected image. Nonparametric correlation analysis was also applied to the stand descriptors on the forest cover map to produce a simplified legend to describe clusters. This procedure however, was impaired by an inadequate distribution of the class data, and was reduced to the use of species groupings and density class.

Classifications resulted in mixed themes due to the heterogeniety of the vegetation, the low reliability of the forest cover map for identification, and the use of algorithms that were designed primarily for the

classification of lower resolution data. The final classification was conservative with respect to the anticipated increase in the number and description of classes reported in the literature. The total gains will be best determined by using actual TM data in an area where more information is available to develop signatures.

ACKNOWLEDGEMENTS

Sincere and deep thanks are extended to Dr. Peter Crown, Flight Commander of the ICARUS Precision Flying Team (IPFT), for his encouragement, assistance, advice, and reviews throughout my graduate program. Dr Crown was also responsible for establishing and obtaining approval for my Interdisciplinary M.Sc. program in conjunction with the Department of Forest Science, a program that was a first in remote sensing.

Appreciation is also extended to Dr. Steve Titus for his encouragement and critical but constructive review of this manuscript. Dr. Titus introduced me to the field of multivariate statistics.

Funding support, direct and indirect, was provided by my employer, the Northern Forest Research Centre (NoFRC) of the Canadian Forestry Service, the Canada Centre for Remote Sensing (CCRS) through its Airborne Remote Sensing Program, and the Northern Affairs Program of DIAND in the NWT. The latter is further acknowledged for its cooperation and interest in this study. In particular, thanks are extended to Jack Gilmore and Chris Carlisle.

Assistance in creating format files was provided by DIPIX Systems Ltd. and were completed with the assistance of the Alberta Remote Sensing Center. Peter Geib and Wendy Power are acknowledged for their periodic assistance. Tom Walsh, NoFRC technician, performed numerous errands on my behalf. His efforts were valuable in assisting me to

maintain progress on a simultaneous, fully committed research program. Cibachrome processing was completed under short time constraints by Jim Baker at NoPRC and his efforts are much appreciated.

Fellow members of IPPT and POP are acknowledged for their support and encouragement, especially since we were all living the daily struggles for survival in the graduate ranks. Acknowledgments would also not be complete without thanking my typist Cindy, a recent M.Sc. graduate herself, for her advice and transcription of my written drafts into what became, finally, a typewritten thesis.

My mother, sister and her husband, and my mother and father-in-law were constantly supportive. Last but certainly not least, my wife Judy deserves the greatest of thanks for her "patience", encouragement and understanding. She was especially tolerant of my many sacrifices of "our" time during the last year and a half of the program, which sometimes seemed like an endless period.

Table of Contents

Chapter		Page
1.	INTRODUCTION	1
	1.1 The Thematic Mapper	3
	1.1.1 Radiometric resolution	4
	1.1.2 Spectral resolution	6
	1.1.3 Spatial resolution	8
	1.2 Study Objectives	8
2.	APPROACH TO THE STUDY	10
	2.1 Study Area	10
	2.2 Data Acquisition	13
	2.2.1 Image data	13
	2.2.2 Ancillary data	18
	2.3 Data Analysis Methodology	19
	2.4 Assumptions and Limitations	25
3.	SIMULATING THE THEMATIC MAPPER	27
	3.1 Radiometric Resolution	27
	3.2 Spectral Resolution	27
	3.3 Spatial Resolution and Geometric Correction	28
	3.3.1 Methods	28
	3.3.2 Results and discussion	30
4.	CORRELATION ANALYSIS	33
	4.1 Rationale	33
	4.2 Methods	34
	4.3 Results and Discussion	39
5.	CLUSTER ANALYSIS	43

5.1 Methods 44

5.1.1 Geographic referencing 44

5.1.2 Cluster analysis procedures 45

5.2 Results and Discussion 52

5.2.1 Geographic referencing 52

5.2.2 Cluster analyses 53

6. UNSUPERVISED CLASSIFICATION 58

6.1 Methods 58

6.1.1 Unsupervised signature generation and classification 58

6.1.2 Post-classification filtering 60

6.2 Results and Discussion 62

6.2.1 Unsupervised signature generation and classification 62

6.2.2 Post-classification filtering 69

7. SUMMARY AND RECOMMENDATIONS 71

7.1 Recommendations 75

7.2 Proposed Role of Satellite Remote Sensing in the NWT 76

BIBLIOGRAPHY 78

APPENDICES 87

LIST OF TABLES

Table — Page

1. Comparison of system parameters for LANDSAT 4 and 5 with LANDSAT 1, 2, and 3.....5
2. NWT forest land area figures.....11
3. Daedalus DS1260 airborne scanner system characteristics.....14
4. Comparison of airborne MSS bands with those of the TM.....14
5. Airborne scanner spatial resolution at specified altitudes.....16
6. Airborne data acquired.....18
7. Portion of DIAND-NWT forest inventory legend system.....35
8. Species groupings for correlation analysis.....37
9. Frequency of class occurence.....38
10. Nonparametric Kendall's rank correlations.....40
11. Partial correlations.....41
12. Affine transformation results for geographic referencing.....53
13. Differences in similarity coefficients with the change from relatively small to large values underlined.....54
14. Number of pixels per cluster.....54
15. Unsupervised clustering analyses performed on the ARIES-II.....61
16. Post-classification filtering employed on the ARIES-II.....63
17. Themes produced from unsupervised analysis number 6.....66

Table Page

18. ARIES-II task TF responses to create format files.....92

LIST OF FIGURES

Figure Page

1. Generalized spectral reflectance curves for vegetation and bare soil.....7
2. Overview flowchart of project components.....20
3. Flowchart of steps in geometric correction.....31
4. Hierarchical clustering directions for an example with four objects.....48
5. Flowchart of cluster analysis sequence.....51
6. Order of spectral bands on tape.....90
7. Image data contained in each record on 11 channel tape.....91

LIST OF PLATES

Plate Page

1. Color composite image of study area.....12

2. Unfiltered and filtered unsupervised results with multiple themes combined (reference Table 17).....67

LIST OF APPENDICES

Appendix — Page

1. Tape format files for 11 channel CCRS JSC airborne MSS tape.....88

2. Control points and residual values from the affine transformation on the uncorrected image....93

3. FORTRAN program to create data files from ARIES-II features files.....95

4. Control points and residual values from the affine transformation on the geometrically-corrected image.....103

5. SPSS version 9 command files for geographic referencing.....105

LIST OF ABBREVIATIONS

AGL	above ground level
CCRS	Canada Centre for Remote Sensing
DIAND	Department of Indian Affiars and Northern Development
FOV	field-of-view
JSC	Johnson Space Center
MDT	mountain daylight time
MSS	Multispectral Scanner
MTS	Michigan Terminal System
NASA	National Aeronautics and Space Administration
NoFRC	Northern Forest Research Centre
NTS	National Topographic System
NWT	Northwest Territories
SE	standard error of estimate
SPSS	Statistical Package for the Social Sciences
TM	Thematic Mapper
UTM	Universal Transverse Mercator

1. INTRODUCTION

Technological advances in higher resolution digital data, have resulted in studies evaluating the utility of these data, in comparison to LANDSAT for incorporation into operational forest inventories (Aldrich, 1979). The use of satellite remote sensing data as an aid to the inventory and management of natural resources has been well documented (McQuillan, 1975; Colwell, 1978; Aldrich, 1979). Little use however, has been made of recent techniques to forestry in the Northwest Territories (NWT), where only generalized data are available for locations and amounts of productive forests' (Bonner, 1982). The results of studies using LANDSAT Multispectral Scanner (MSS) digital data for broad forest and land cover classification, though useful in their own context, have varied in the number of spectral classes and classification accuracy (Beaubien, 1979; Lee, 1980; Kalensky *et al.*, 1981; Pala and Jano, 1981).

This variation has been due to three general areas of difficulty in forest classification. One difficulty was in deciding on the best approach for image classification since the forester can choose among many methods. Townshend and Justice (1981) concluded from their review, that there was little evidence in the literature to guide users in choosing between the many techniques. Decisions have to be made on digital preprocessing, and between unsupervised clustering

---

'Productive forest land is land that is capable of producing a merchantable stand within a reasonable length of time(Bonner, 1982).

and supervised methods in image classification, for example, and between the hybrid techniques that have been developed (Hoffer and Fleming, 1978).

A second problem was in identifying how remote sensing should be incorporated into an inventory program, with consideration to existing information. Helms and Shain (1981) encountered several problems in incorporating satellite image analysis into an existing forest inventory program including:

1. the acquisition of data during the desired time frame;
2. the specialized image analysis equipment and staff expertise requirements;
3. the compatibility of satellite-based vegetative classifications with ground observations; and
4. the difficulties in conducting accuracy analyses due to resolution differences of conventional inventory methods.

Separate consideration to levels of information are therefore needed to facilitate incorporation of satellite remote sensing into inventory programs. This can be achieved for example, by devising a classification legend to use with remotely sensed data, that would also be compatible with the more detailed classification legend often used at larger scales.

The third and perhaps most important problem in applying image classification was that the forest and wildlands exhibit variability in spectral responses due to

variability in species composition, stand density, age, percent slope, and aspect (Beaubien, 1979). This variability was not accounted for in a previous in-house LANDSAT digital classification project at the Northern Forest Research Centre (NoFRC) funded by the Department of Indian Affairs and Northern Development (DIAND). The spectral classes were not compatible with the NWT forest inventory legend in terms of numbers and descriptions of types, and were generally not of sufficient detail. This problem was anticipated to be resolved in part by using higher resolution digital data.

During the late 1970's and early 1980's, plans were made by the National Aeronautics and Space Administration (NASA) to launch a new generation of resource satellites which would have onboard, a new MSS called the Thematic Mapper (TM). The TM was designed to acquire data at an estimated 10 times the resolution of previous MSS systems (Ahern *et al.*,1980). Since actual TM data were not available, a simulation study using airborne scanner data was suggested to address the three areas of forest classification difficulties. The study reported here describes the results of such a simulation.

1.1 The Thematic Mapper

For this study, knowledge of TM data characteristics with some indication of expected performance was useful in defining specifications for obtaining airborne data, and in performing classifications. TM data may be discussed in

terms of radiometric, spectral and spatial resolution characteristics. An understanding of Thematic Mapper sensor characteristics can be facilitated by comparison to the more familiar LANDSAT MSS. In addition, practical considerations for classifying the higher resolution data can be obtained by reviewing pertinent simulation studies.

The initiation of higher resolution resource satellites began with the launch of LANDSAT 4 on July 16, 1982, and was followed by LANDSAT 5 on March 1, 1984. Numerous papers have been published which describe, in varying detail, the differences between these sensors (Table 1) (Mcquillan, 1978; Covault 1982; Smith, 1982). During this study the TM was classified as an experimental research sensor. TM data are now available for only a few selected areas.

1.1.1 Radiometric resolution

Radiometric resolution refers to the number of quantization levels in the data received. The TM provides for 8 bit ($2^8 = 255$) data whereas the LANDSAT MSS provides for 6 bit ($2^6 = 64$) data. The greater radiometric sensitivity of the TM over the MSS should permit detection of smaller changes in radiometric magnitudes within a given spectral band, and provide for greater sensitivity to relationships between spectral bands (Freden and Gordon, 1983). Tucker (1980) reported classification improvements for bands 3 and 4 based on simulated TM data to be 1 to 3% more accurate than the MSS.

Table 1: Comparison of system parameters for LANDSAT 4 and 5 with LANDSAT 1, 2 and 3.

System Parameters		LANDSAT 4 and 5		LANDSAT 1,2,3
		Thematic Mapper	Multispectral Scanner	Multispectral Scanner
Radiometric resolution(bits/pixel)		8	6	6
Quantization levels		256	64	64
Spectral bands (wavelength-um)	1	0.45- 0.52	0.5-0.6	0.5-0.6
	2	0.52- 0.60	0.6-0.7	0.6-0.7
	3	0.63- 0.69	0.7-0.8	0.7-0.8
	4	0.76- 0.90	0.8-1.1	0.8-1.1
	5	1.55- 1.75		
	6	10.40-12.50		
	7	2.08- 2.35		
Nominal pixel spatial resolution (m)		bands 1-6: 30 band 7: 120	82	79
Detectors per band		bands 1-5,7: 16 band 6: 4	6	6
Swath width (km)		185	185	185
Data rate (Mbps)		85	15	15
Cross track scanning		bidirectional	west-east	west-east
Approximate number of pixels per scene(million)		300	32	32
Nominal altitude (km)		705	705	913
Repeat cycle (days)		16	16	18
Launch dates		LANDSAT 4:July 16, 1982 LANDSAT 5:March 1, 1984		LANDSAT 1:July 23, 1972 LANDSAT 2:January 22, 1975 LANDSAT 3:March 5, 1978

1.1.2 Spectral resolution

Spectral resolution refers to the width and number of spectral bands with improved resolution associated with narrower and more numerous bands (Table 1). The placement of spectral bands can be visualized by comparing Figure 1 and Table 1.

Based on simulation studies, TM channels 1, 2, 3 and 4 were considered to be well placed for sensing vegetation, and substantial improvements in classification performance were expected from the increase in spectral resolution alone (Tucker, 1978). Few studies however, have reported on the use of the TM thermal infrared band 6[2]. The main difficulty was the combining of emitted and reflected data in frequently used classification algorithms. The possible classification results using all 7 channels could not be evaluated since most efforts have been with reflected radiation bands (Tucker, 1978). Latty and Hoffer (1981) concluded that the greater spectral resolution will not necessarily result in greater classification accuracies. They further stated that results were highly data and application dependent, and would generally not improve when using more than 4 channels.

---

[2]pers.communications Dr.Roger Hoffer, Purdue University, West Lafayette, Indiana. June, 1981.

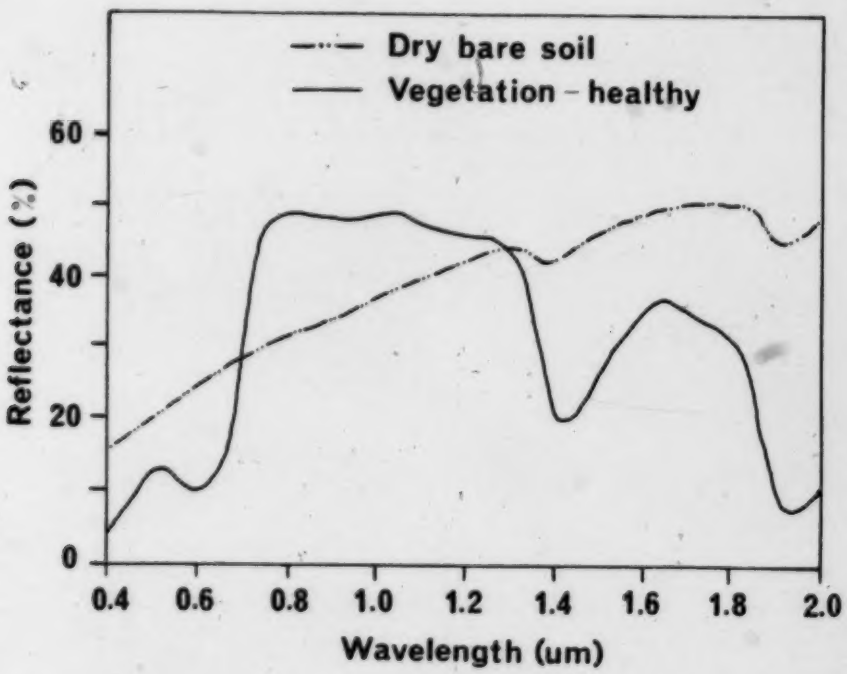

Figure 1: Generalized spectral reflectance curves for vegetation and bare soil.

Source: Adapted from Lillesand and Kiefer, 1979.

1.1.3 Spatial resolution[3]

The TM has a 30 m pixel cell compared to the 79 m for the MSS for all bands except for band 6 which is 120 m (Table 1). Theoretically, this allows areas as small as 2.5 to 4 ha (assuming a minimum of 28 contiguous pixels as a minimum for one theme) to be classified (U.S. Geological Survey, 1982). In evaluating the utility of smaller ground resolution cells, Hyde *et al.*, (1981) concluded that though mixed pixels will still occur, area classification and overall classification accuracy should generally improve although data processing time will be greater. The improved spatial resolution connotes an increase in the number of pixels to blanket a land cover type of a given size (Hyde *et al.*,1983). The benefit of a 30 m pixel cell would therefore be in obtaining information on smaller spectrally variable areas. In practical terms however, though smaller and more complex areas may be classified, the actual accuracies should depend on the resource being classified.

1.2 Study Objectives

The main objective of this study was to develop a methodology to classify simulated Thematic Mapper data for broad forest and land cover types. The classification exercise also served as the vehicle to address the difficulties in forest classification.

---

[3]spatial resolution for the purposes of this study is considered to be pixel size. Further discussion is available in Forshaw *et al.*, (1983).

Four subobjectives were established to achieve the main study objective. The first subobjective was directed toward data manipulation:

1. To simulate the TM spatial resolution, model the image distortions and geometrically correct the airborne data.

The direction for the study was reflected in the remaining subobjectives which involved aspects of methodology development:

2. To rationalize the unsupervised approach taken in this study and consequently address the first difficulty in forest classification.

3. To relate image data to existing map information, and to account for the major components of vegetative reflectance. This objective also addressed the second and third difficulties in forest classification by developing a simplified legend system to describe clusters from statistically-based image classifications.

4. To determine if cluster analysis could be used as an aid in determining the number and possible content of output clusters. One difficulty with the unsupervised approach was in determining the final number of spectral classes. The results of such analysis was proposed as a possible tool for subsequent unsupervised signature generation and classification.

2. APPROACH TO THE STUDY

This chapter begins with a description of the study area, followed by the data acquired, data analysis methodology, and assumptions and limitations.

2.1 Study Area

The NWT has a unique information need and spectrally complex area for classification. The information need is due to incomplete quantitative forest survey coverage (Bickerstaff *et al.*, 1981). The NWT has a larger land area than any of the Provinces, and the most accurate figures available (Bonner, 1982) estimate 19% of the total land base of 3.246 million square km is in forest land. Not all of the forest land is commercially productive however, and published figures for the total productive forest land have varied (Table 2). Figures for the specific locations and amounts are unknown. Current forest values are low due to low utilization mainly for local use. Increasing demands for forest products however, are creating a need for better managed forests (Indian and Northern Affairs Canada, 1983).

The study was conducted in a 10 by 14 km area in the immediate vicinity of Fort Smith, NWT (Plate 1). The actual study area in the southwest corner of the 1:250 000 National Topographic System (NTS) Map 75D, has a heterogeneous mixture of pure and mixedwood stands with frequent occurrences of treed bog and bog wetlands. This area is in the Northwestern Transition Section (B.27) of the Boreal

Table 2: NWT forest land area figures.

Land Type	Forest Area (1 000 square km)		
	Bowen 1978	Bickerstaff et al. 1981	Bonner 1982
Forest land	307		615
Productive forest land		151	143
Wildland	2 939	1 514	2 488
Total	3 246	1 665	3 246

Forest Region (Rowe, 1972). Black spruce (*Picea mariana*(Mill.)B.S.P.) is most abundant, with white spruce (*Picea glauca*(Moench) Voss), jack pine (*Pinus banksiana* Lamb.), trembling aspen (*Populus tremuloides* Michx.), and balsam poplar (*Populus balsamifera* L.) also occurring. Tree species identification in the North is frequently difficult due to the presence of hybrids. Fire is a natural ecological factor, and wildfires have contributed greatly to the variations in species and age of vegetation (Indian and Northern Affairs Canada, 1983). The extensive fire history and variable growth conditions, typical of many wildland areas in the NWT, add to the spectral variability within the area.

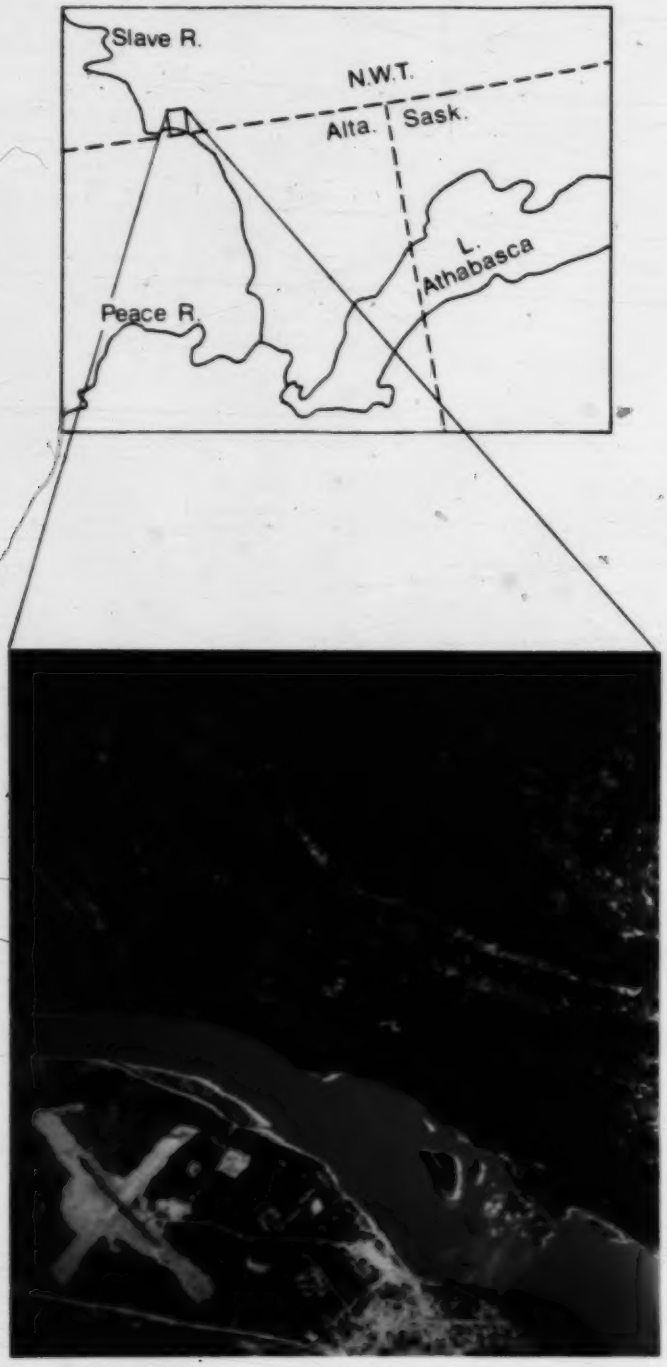

Plate 1: Color composite image of the study area with the Ft. Smith airport in lower left corner. (distance across the image is 10 km)

2.2 Data Acquisition

The data acquired included the airborne scanner data, aerial photographs, and a forest cover map of the study area.

2.2.1 Image data

The Canada Centre for Remote Sensing (CCRS) Daedalus DS 1260 11-channel airborne multispectral scanner (Zwick *et al.*,1980) was used in this study (Table 3). This scanner has been used in several TM simulation projects (Ahern *et al.*,1980; Derenyi and Yazdani, 1981; Teillet *et al.*,1981; Tomlins, 1981). The DS 1260 is a rotating mirror scanner with a 85.9° swath (73.7° with S-bend correction for roll and mirror scan velocity[4]). Data are acquired on an 8-bit radiometric scale and the radiometric resolution is the same as for the TM.

TM spectral resolution for this study was simulated by selecting those airborne channels which most closely corresponded to the TM data in terms of band location and width (Table 4). This follows the procedure of previous projects (Ahern *et al.*,1980; Teillet *et al.*,1981).

There were several important spatial considerations to using an airborne MSS, to simulate data to be obtained from a satellite 705 km in space. The airborne MSS was mounted in the Falcon Fan Jet which has an aircraft ceiling of 11 000 m

---

[4]H.R. Edel.CCRS processing of airborne data. unpublished, undated internal report. Canada Centre for Remote Sensing. p.8

Table 3: Daedalus DS1260 airborne scanner system characteristics(excerpt from Zwick *et al.*,1980).

System Characteristic	
radiometric resolution	8 bits/pixel
spectral resolution	11 channels(Table 4)
spatial resolution	2.5 mrad
scan angle	73.7°(with S-bend correction)
number of pixel per scan line	716
scan rates	12.5,25,50,100 scans/sec

Table 4: Comparison of airborne MSS bands with those of the TM.

SPECTRAL CHANNELS			
AIRBORNE MSS		THEMATIC MAPPER	
Channel No.	Wavelength (nm)	Channel No.	Wavelength (nm)
1	380- 420		
2	420- 450		
3	450- 500	1	450- 520
4	500- 550		
5	550- 600	2	520- 600
6	600- 650		
7	650- 690	3	630- 690
8	700- 790		
9	800- 890	4	760- 900
10	920-1100		
11	1550-1750	5	1550- 1750
		6	10400-12500
		7	2080- 2350

above sea level (Airborne Operations Section, 1977). The above ground altitude would vary with the ground elevation and govern the spatial resolution of the data acquired (Table 5). These figures were determined using the equation:

$$\text{height (m above ground level(AGL))} = \frac{\text{spatial resolution(m)}}{2.5 \times 10^{3} \text{ radians}}$$

TM spatial resolution was simulated by flying as close to the aircraft ceiling as possible, followed by pixel resampling.

There were illumination problems to contend with due to angle of illumination and geometric distortions (sensor angle of view) (Crown and Pawluk, 1973), and atmospheric effects. The latter occurs in the atmosphere between sensor and target, and between the radiation source and the target (Silva, 1978). Though there is simultaneous scattering and absorption of target radiation, it will vary depending on variation in altitude and atmospheric conditions (Horvath *et al.*, 1970). Considering that these effects will also influence satellite data, airborne scanner data obtained at the highest elevation as possible, would seem to best simulate spectral information to be received by the TM.

Airborne MSS scan rates are operator selectable and dependent on aircraft speed and altitude (Zwick *et al.*, 1980). The slowest scan rate (12.5 scans/sec) was used in this study. The S-bend correction is available at the

Table 5: Airborne scanner spatial resolution at specified altitudes.

Spatial Resolution(m)	Altitude(m AGL)
2.5	1000
10	4000
20	8000
25	10000
30	12000

slowest scan rate[3] which would be used at higher altitudes.

Geometric distortions are due to aircraft altitude and changes in ground resolution cell size with greater scan angle (Kindelan *et al.*,1981). Geometric effects due to varying aircraft altitude especially at low altitudes, are much greater than for the satellite which has a relatively more stable platform (Kindelan *et al.*,1981). Pertaining to scan angles, the airborne MSS is 73.7°, whereas for LANDSAT 2 it is 5.8° (U.S. Geological Survey, 1979), and for LANDSAT 4 it is 7.5°. It was not possible to limit airborne scan angles to the same magnitudes as the LANDSAT satellites because of insufficient pixels for analysis. To reduce geometric and atmospheric distortions, most studies therefore work within a field-of-view (FOV) of 20° either side of the nadir (Zwick *et al.*,1980). The swath can consequently be calculated to determine the width of ground covered[4].

---

[3] ibid.,p.10.
[4] personal communications. internal memo from Dr. G. Tomlins

$$\text{Swath(m)} = 2 \times \text{altitude(m)} \times \tan[\text{FOV(scan angle)}]$$

There was one other consideration which pertained to aircraft direction when obtaining data. For more equal illumination across the swath and to prevent "hot spots", the aircraft should fly either into or out of the sun.

In summary, there are important considerations in using the airborne scanner for acquiring simulated TM data. These considerations confirm two concerns published in Audet and Thomson (1981);

1. simulated data should be as representative of the TM as possible, and
2. the effects of resampling high resolution data to simulate TM data should be observed.

The scanner should therefore be flown in a specified direction at the highest altitude possible, with only data within a limited FOV being used in analysis.

All aircraft data were acquired by CCRS, one year later than first planned due to forest fire, smoke, and aircraft problems. The airborne data acquired are summarized in Table 6. Data acquired for the TM simulation using the scanner and the 1:66 000 scale photography were acquired simultaneously. An approximate pixel size of 25 m was acquired. To reduce effects caused by the atmosphere, only data within ±20° of the nadir were used for analysis.

'(cont'd) to C. Goodfellow, Canada Centre for Remote Sensing. June, 1980.

Table 6: Airborne data acquired.

Study Area: Fort Smith, NWT

Date and Time: July 22, 1982, approx. 15:00 hrs MDT

Altitudes: 10 060 m, 4000 m AGL

Direction: into the sun

CCRS task no.: 82-34

Sensors: 1. DS 1260 airborne scanner
(ref. Table 4 for spectral wavelengths)
S-bend correction
CCRS - JSC CCT no. AR0596

2. Two RC-10 aerial cameras, sun angle >35°,
65% forward overlap, focal length=152 mm,
nominal scales: 1:66 000, 1:26 200

a) 1 camera with Kodak 2443 color infrared film, catalogued under A37848

b) 1 camera with Kodak 2445 color negative film, catalogued under A37849

2.2.2 Ancillary data

A 1:25 000 scale forest cover map was supplied by forestry staff of DIAND in Fort Smith. Forest cover maps are generally completed under contract and verified by forestry staff. Some ground checking verified DIAND observations that the type map was highly variable in quality and reliability, and that many of the stand boundaries were poorly represented'. The acquired photography in combination with the type map was used for the spectral classes.

' personal communications. memo from C.R. Carlisle, Forest Resources-NWT, DIAND, Fort Smith, NWT. December, 1981.

The airborne data acquired was analyzed using the DIPIX Systems Limited ARIES-II image analysis system (ARIES-II) at the Department of Soil Science, University of Alberta, and at the Alberta Remote Sensing Center, Alberta Environment, Edmonton.

2.3 Data Analysis Methodology

The project proceeded through several distinct and separate phases which are summarized in a flowchart (Figure 2). Geometric correction required development of extra procedures for data manipulation and are described in the next chapter. This section begins with the rationale for selecting the image classification technique employed and follows with the actual methodology.

There are two approaches to image analysis applicable to forestry. The first approach is based on knowledge of the spectral characteristics of the resource being classified, by concentrating on spectral signatures and radiance values in the various wavelength bands. The second approach is purely statistical since the use of many algorithms (e.g., maximum likelihood, unsupervised clustering, principal components), can be viewed as a statistical manipulation of the digital data. The approach taken in this study was a combination of both whereby statistical methods were used to assist image analyses.

The procedures employed started with a review of completed projects to summarize experience gained from other

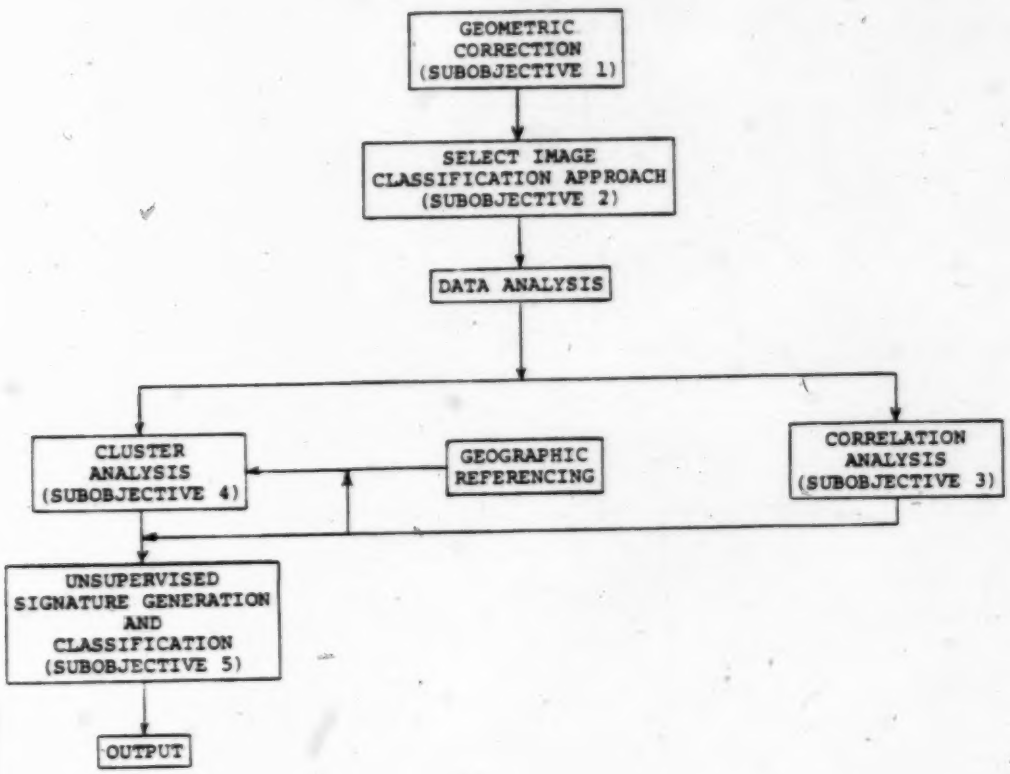

Figure 2: Overview flowchart of project components.

simulation studies, and to ensure that this study would not be a repeat of work previously undertaken. Five projects using simulated TM data were reviewed beginning with Williams and Stauffer (1980), who concluded that the major challenge would be in processing much larger data volumes than with the LANDSAT MSS.

Lee (1981) obtained 16.6 m airborne data and resampled the data to 30 m. Identical training areas were created which included cutover, young growth, poor site mature forest, medium site mature forest, good site mature forest, regeneration, water and hardwoods. No further descriptions were given as to the structure and content of the land cover classes. Lee (1981) stated that the number of unclassified pixels depended on the number of channels used and on the pixel resolution. This was an example of the effects from greater degrees of collinearity due to the use of more channels. This trend was also found by Thiruvengadachari (1981) who obtained 15.7% unclassified pixels with 11 channels compared to 1.14% with 4 channels. Greater pixel resolution also increases the number of unclassified pixels although the percentage of the scene unclassified may decrease. Homogeneous areas at smaller spatial resolutions exhibited greater variability at larger spatial resolutions. These results were closely related to concerns by Crist (1984) in using TM data for training samples with traditional supervised classification approaches.

Teillet *et al.*, (1981) compared classification performance using simulated TM and LANDSAT MSS data in terms of specified land cover classes. Their study area was in a mountainous area in British Columbia in which several softwood species occurred in pure or mixed stands with occasional stands of hardwoods. Mountain pine beetle (*Dendroctonus ponderosae* Hopk.) and western spruce budworm (*Choristoneura occidentalis* Free.) infestations existed in the study area. Selective logging occurred on lower slopes and had the main effects in altering stand composition and density. In addition, apart from species, age and density were also major factors influencing softwood class distribution (Beaubien, 1979). The above factors compounded with vegetation reflectance differences in relation to the position of vegetation on the slope, resulted in a potential for large spectral variations to contend with in classification. Cover type maps were reported to be available though only species information appeared to have been used. Results were compared for two sets of land cover classes using supervised classification and feature selection based on average transformed divergence (Goodenough *et al.*, 1978). Classification accuracy improved with simulated TM data by reducing misclassification error from 33% to 17%. General inventory was expected to require all spectral information (Teillet *et al.*, 1981). This conclusion was in slight disagreement with Lee (1981) who warned that using all channels would increase the number of

unclassified pixels.

Dean and Hoffer (1982) used the supervised approach and compared the effectiveness of feature selection using principal components and transformed divergence with different classification algorithms, and 3 and 4 channel or component subsets. A 4 channel subset was concluded to generally perform more accurately than principal components or transformed divergence.

Study objectives by Nelson *et al.*, (1984) included determining the most useful TM bands for differentiating boreal forest cover types and insect damaged from undamaged forest, providing an assessment of classification accuracy, and identifying land cover types subject to confusion. In agreement with Beaubien (1979), Nelson *et al.*(1984) recognized that accurate forest classifications were often difficult especially since cover type changes tend to be gradual. Discriminant analysis was used for band selection based on land cover classes as interpreted from aerial photographs. Results included verification of a trend by Swain and Davis (1978) that decreasing accuracy may be a function of a fixed training sample size and increasing dimensionality.

The forestry TM simulation papers provide evidence that the supervised approach predominates in the literature. This approach is suitable when ancillary and ground information are available, which is not always true in the NWT. Assumptions in the supervised approach are that classes are

spectrally separable from other classes, and have informational value (Landgrebe, 1973). A difficulty, however, may be in the selection of training sites because of the variability of the spectral response (Townshend and Justice, 1980). Cover types may not be spectrally homogeneous, and samples chosen may not truly represent classes of interest (Goldberg *et al.*, 1975; Mroczynski *et al.*, 1980), nor account for all spectral variations (Hoffer and Fleming, 1978; Hoffer and Swain, 1980). One reason for inadequately selected samples may be insufficient consideration of factors influencing canopy reflectance which can be helpful in identifying and grouping vegetation classes (Kachhwaha, 1983). Colwell (1974) and Beaubien (1979) provided discussions of factors affecting vegetative cover reflectance, and some accounting of these factors were attempted in this study, and is described in the correlation analysis chapter.

In a wildland situation, such as in the NWT, it cannot be assumed that the image analyst has prior knowledge of the structure and identity of land cover classes, and that they are spectrally separable (Nelson and Hoffer, 1979). Unsupervised approaches are therefore necessary for vegetation mapping in northern regions (Beaubien, 1979) when there is no prior information on class structure (subobjective 2). The difficulty in the unsupervised approach however, is that the analyst must specify the number of clusters (classes) that the computer should

generate. The decision on the number of classes has significant implications. A small number of classes may result in multimodal spectral classes with large variances, whereas too many classes will result in problems in identifying the output classification (Nelson and Hoffer, 1979). In addition, little knowledge on complex heterogeneous areas can make the assignment of information labels to spectral classes difficult. The premise of this study was that use of a cluster analysis statistical package (subobjective 4) on a mainframe computer, would facilitate more detailed analysis of the number of categories to be specified, when using the unsupervised method on the ARIES-II. Also, a system for geographic location of pixels would provide a means for placing information labels on the clusters, where vegetative cover information was available (geographic referencing, Figure 3).

2.4 Assumptions and Limitations

There were several assumptions and limitations to the study:

1. Beyond corrections from ground checking, the forest cover map and the aerial photographs would be considered adequate for placing labels on clusters. This was the most significant assumption especially given the common practise of accepting existing maps as correct without challenging the information content (Krebs, 1982). Despite the poor planimetric accuracy of the forest

cover map, it had to be used in geometric correction since no other map of suitable scale was available. The forest cover map was also used in the correlation analysis and geographic referencing portions of the data analysis.

2. Airborne MSS data were assumed to be a reasonable simulation of the TM reflectance data.
3. Airborne scanner data were acquired at a time of day different from the TM overpass. Atmospheric and illumination effects therefore may or may not be the same as for TM data.
4. A possible limitation in classification is the use of data from a single date. Multidate data may improve forest classification results since they account for phenological changes (Kalensky and Scherk, 1975; Purvee and Braun, 1978; Long and Hopkins, 1984).

3. SIMULATING THE THEMATIC MAPPER

The purpose of this chapter is to summarize the airborne data acquisition and preprocessing that were performed for the TM simulation.

3.1 Radiometric Resolution

The image data acquired by the scanner required no manipulation since it was on the same 8-bit radiometric scale as the TM.

3.2 Spectral Resolution

TM spectral resolution was to be simulated by selecting the appropriate channels from a 11-channel tape ordered from CCRS. This tape was created using the Johnson Space Centre (JSC) Universal band interleaved format (Murphy, 1979). Four format files were created to read 3 of the 11 airborne MSS channels at a time and are more fully described in Appendix 1. Only channels 5, 7 and 9 (Table 4) were used for analyses. Channel 3 was unsuitable because of haze, channel 6 because of line drop-outs, and channel 11 data were intermittent. The loss of the latter was especially unfortunate as this reflected infrared region has not previously been acquired by the LANDSAT MSS.

3.3 Spatial Resolution and Geometric Correction

Geometric correction and resampling of the airborne data were necessary to simulate the TM spatial resolution and to facilitate tabulating classifications by 10- or 1- km Universal Transverse Mercator (UTM) cell. Standard procedures were first proposed for geometric correction. These included the location of control points on the NTS map and displayed image, and the performance of an affine transformation. The latter would be followed by pixel resampling. Correction of airborne line scan images can be difficult however, due to platform instability, systematic scanning distortions, viewing perspective and surface topography (Leckie, 1980). Control points in northern regions can also be difficult to locate (Schreier *et al.*, 1982). These factors placed constraints on the methods that were employed.

3.3.1 Methods

The standard procedure would be to locate control points on a displayed, uncorrected image and corresponding 1:50 000 NTS map. Published 1:50 000 NTS map sheets are rare for the NWT, and none existed for the study area. The largest scale map available was the 1:25 000 forest cover map which did not have UTM grid lines. The 10-km UTM grid lines from the 1:250 000 NTS map was therefore triangulated onto the 1:25 000 forest cover map. The positions of the 1-km UTM grid lines were subsequently interpolated.

Potential control points on the displayed digital image were frequently not found on the map but were found on the photography. A photogrammetric radial line triangulation procedure was therefore employed to transfer desired control points from the aerial photograph to the map. UTM coordinates for the control points were subsequently derived. An assumption was made that the 1:66 000 scale photographs were approximately vertical to permit application of the radial line triangulation principle. The principle is that in vertical photographs, angles about the nadir are equivalent to angles measured on the ground, regardless of general scale change and amount of relief (Hall, 1978). Since a map is an orthographic projection on which every point is at the same scale, angles measured will be the same as on the ground.

Following control point location, a first order affine transformation was undertaken followed by cubic convolution resampling (Moik, 1980). The affine transformation is a least squares multiple regression with two linear terms, and is of the form (Kirby and Steiner, 1978):

$$NORTHING = B + B_1 PIXEL + B_2 LINE$$

$$EASTING = B + B_1 PIXEL + B_2 LINE$$

An affine transformation accounts for rotation, displacement, scaling and skew (Moik, 1980), and is evaluated by studying statistical results (e.g., standard error of estimates) and analyzing residuals. The statistical values were provided in a table as an output product of the

polynomial transformation task.

The spatial resolution of the airborne data was 25 m (section 2.2). Resampling was therefore necessary to simulate the TM spatial resolution of 30 m. Resampling was accomplished using cubic convolution, an algorithm most commonly used in vegetation applications (Kalensky *et al.*, 1981; Ueno, 1982), and generally providing finer image detail than other resampling algorithms (Moik, 1980). The sequence of steps undertaken in geometric correction are summarized in a flowchart (Figure 3).

3.3.2 Results and discussion

Ten control points were located and standard errors of 63.0 and 57.1 m for the easting and northing estimates, respectively, were obtained. A table summarizing the locations and residual values of the individual control points is given in Appendix 2. The accuracy of a transformation is directly dependent on the accuracy of the control points and will vary from pixel to pixel (Davis and Kenue, 1977). The errors obtained were a nominal two pixels and greater than the accuracy criterion established at one pixel or less. Since geometric correction and resampling were needed however, the errors had to be accepted and were considered reasonable within the methodology that had to be employed. The magnitude of errors were dependent upon the number of control points, their location error, and their spatial distribution (Shlien, 1979; Orte, 1981). A concern

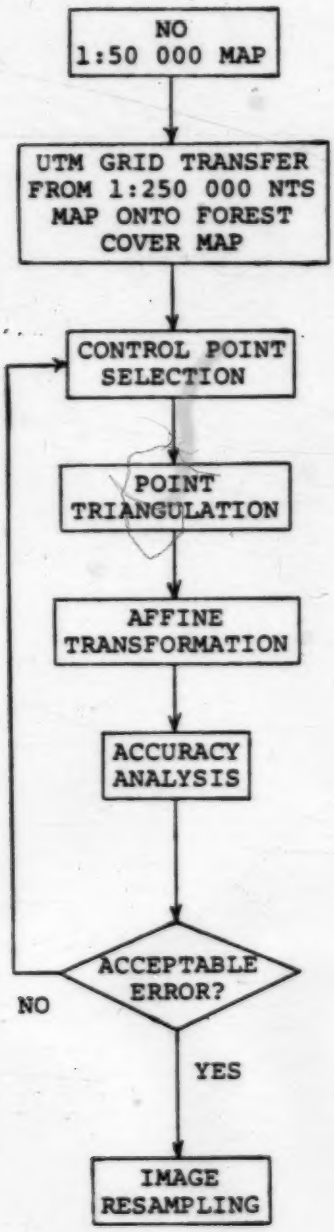

Figure 3: Flowchart of steps in geometric correction.

might be raised as to the adequacy of the 10 control points used and their distribution. Sources of location error were typical of those occurring in the geometric correction process and included map inaccuracy, map measurement error, and image measurement error (Haymond, 1982). The above sources of error were possibly the largest due to the poor planimetric accuracy of the forest cover map which greatly affected the accuracy of triangulated points. These errors may have been smaller if a 1:50 000 NTS map sheet had been available. Cubic convolution resampling to the TM 30 m resulted in a study area configuration of 450 lines by 320 pixels.

4. CORRELATION ANALYSIS

The purpose of the correlation analysis was to determine the extent to which the forest cover map and aerial photographs could be used to describe spectral classes from classification. The approach used was to develop a simplified legend system for placing information labels on classified groupings.

4.1 Rationale

A problem often reported with the unsupervised approach has been the assignment of meaningful information labels to spectral classes (Mayer and Fox, 1981). Relatively more effort unfortunately, has been devoted to classifying image data than in presenting the information in a form that a user might easily understand. The problem may be due to an incomplete understanding of the relationship between computer derived statistical classes and actual cover classes (Merchant, 1981).

Remote sensing digital data are in the form of relative radiance values in a feature vector for each pixel, the latter representing a relatively large ground area. The relationship between the radiance values and the legend system must therefore be established to describe forest stands on the ground. Forest classification is based on a set of variables that are recognized on aerial photographs (Hilborn, 1981) and in the field. These variables often include species composition, stand density, age, height and

others such as site index (Hegyi and Quenet, 1981). In combination, these variables are also the factors which will largely govern the reflectance from forest stands. The approach in this study was to simplify the "open" forest inventory legend system (i.e., all combinations of species and height, age, and density class permitted) used by DIAND in the NWT (Table 7). The reduced legend would be a compatible subset of the full legend, and could be used as information labels in ensuing classifications.

4.2 Methods

Correlation analyses were performed once all forest cover types in the study area were tabulated from the forest cover map for species composition and stand attributes (Table 7). Beaubien (1979) reported that the younger and/or denser a stand, the greater its overall reflectance. If significant correlations could be found with height and age, and age and density, a simplified legend system could then be developed consisting only of species composition and density. In addition, although age is an important contributor to spectral reflectance, it is not possible to determine age class alone from radiance values since other factors are always involved. The third subobjective and a null hypothesis was established with the relation to be determined from the statistical results of two secondary hypotheses:

Table 7: Portion of DIAND-NWT forest inventory legend system.

Species	
Sw	white spruce
S	undifferentiated spruce
Sb	black spruce
L	larch(*Larix laricina* (Du Roi) K.Koch)
P	pine (mainly jack or hybrids)
A	aspen
Po	balsam poplar
B	birch (*Betula papyrifera* Marsh.)
H	mixed hardwoods

Stand Attribute	Class Number
Density class (crown closure %)	
1 - 20	1
21 - 40	3
41 - 60	5
61 - 80	7
81+	9
Age class (years)	
1 - 20	1
21 - 40	3
41 - 60	5
61 - 80	7
81 - 120	10
121 - 160	14
161 +	18
Height class (m)	
.01 - 2.50	0
2.51 - 7.50	1
7.51 - 12.50	2
12.51 - 17.50	3
17.51 - 22.50	4
22.51 - 27.50	5
27.51 - 32.50	6
32.51 +	7

Subobjective 3: To relate image data to existing map and air photo information, and to account for the major components of vegetative reflectance.

Hypothesis: A simplified forest and land cover classification system cannot be developed to be compatible with the open legend scheme employed in the NWT.

Subhypothesis A: Stand height does not vary with stand age.

Subhypothesis B: Density does not vary with the age of the stand.

Since the forest cover types were class data, the distribution could not be assumed Gaussian. Non-parametric Kendall's coefficient of rank correlation was therefore performed using a statistical significance level of 0.20. Kendall's was chosen over the more commonly used Spearman's method. Kendall's is more exact and approaches normality with increase in sample size more quickly than Spearman's (Kieth and Cooper, 1974). It may therefore be a more effective technique to use. A major weakness of rank correlation statistics however, is the occurrence of tied ranks. The Statistical Package for the Social Sciences (SPSS) (Nie *et al.*, 1975) employs a correction to the formulation and was therefore used.

The forest cover types were tabulated according to a major species (e.g., spruce, pine, aspen) and species mix (e.g., spruce/aspen, pine/aspen) to constitute the first analysis set (Table 8). The species mix was also joined with the appropriate species to increase the sample size and to

Table 8: Species groupings for correlation analyses.

Label	Species	Sample size
Analysis set 1		
S	white spruce	20
SA	white spruce/aspen	23
P	pine	13
PA	pine/aspen	10
Sb	black spruce	25
SbA	black spruce/aspen	6
A	aspen	20
AConifer	aspen/conifer	32
Analysis set 2		
S, SA	white spruce, white spruce/aspen	43
P, PA	pine, pine/aspen	23
Sb, SbA	black spruce, black spruce/aspen	31
A, AConifer	aspen,aspen/conifer	52

form the second analysis set (Table 8). Partial correlations assuming normality were also calculated using SPSS to determine if there were any differences on the magnitudes and significance levels of the correlations when the other variable, not being correlated, was held constant.

All cover types in the study area were also tabulated according to frequency of height, age, and density class occurrence (Table 9). These figures provided an indication of the range and identified particular classes which occurred more frequently than others.

Table 9: Frequency of class occurrence.

Label	Height class	Height number	Age class	Age number	Density class	Density number
Analysis set 1						
S	1	7	5	6	1	1
	2	8	7	10	3	7
	3	4	10	4	5	10
	4	1			7	2
SA	1	3	5	3	1	3
	2	10	7	19	3	7
	3	7	10	1	5	13
	4	3				
P	0	1	5	1	5	8
	2	3	7	11	7	5
	3	8	10	1		
	4	1				
PA	1	3	3	3	5	8
	2	3	5	2	7	2
	3	2	7	5		
	4	2				
Sb	0	1	3	2	3	5
	1	22	5	2	5	11
	2	2	7	8	7	9
			10	13		
SbA	1	5	5	3	1	1
	2	1	7	2	3	4
			10	1	5	1
A	1	1	3	2	1	3
	2	11	5	3	3	6
	3	6	7	14	5	10
	5	2	10	1	7	1
AC	1	4	3	2	1	2
	2	7	5	4	3	8
	3	19	7	24	5	21
	4	1	10	2	7	1
	7	1				
Analysis set 2						
S,SA	1	10	5	9	1	4
	2	18	7	29	3	14
	3	11	10	5	5	23
	4	4			7	2
P,PA	0	1	3	3	5	16
	1	3	5	3	7	7
	2	6	7	16		
	3	10	10	1		
	4	3				
Sb,SbA	0	1	3	2	1	1
	1	27	5	5	3	9
	2	3	7	10	5	12
			10	14	7	9
A,AC	1	5	3	4	1	5
	2	18	5	7	3	14
	3	25	7	38	5	31
	4	1	10	3	7	2
	5	2				
	7	1				

4.3 Results and Discussion

Despite some low magnitudes, there were generally significant correlations between height and age (Table 10, 11). Though not consistently significant across all species groupings, hypothesis A was rejected with the conclusion that stand height does vary with stand age in at least, most cases. A reason for some of the low correlations was large numbers of a particular height or age class, which would limit the ranges of class data (Table 9).

There was no consistant statistical relationship between density and age (Table 10), even when the variation in height was controlled (Table 11). Since there were a few significant correlations however, subhypothesis B was rejected with the conclusion that stand density does vary with stand age, in at least a few instances. A review of the individual cover types and Table 9 resulted in two possible reasons for the very few significant correlations. One reason was no apparent pattern of relationship among density or age classes and therefore, no correlation. There was for example, some very old black spruce and black spruce/aspen types of varying densities. This would result in difficulties correlating a particular density with a particular age. The second reason was relatively large numbers of certain age or density classes relative to the total for a species group. The result was an insufficient number of classes to cover a wide enough range to produce a correlation. Other possible reasons for the results obtained

Table 10: Nonparametric Kendall's rank correlations.

Label	Correlation Coefficients height/age	density/age
Analysis set 1		
S	0.572 *	0.016
SA	-0.109	0.183
P	0.578 *	0.0
PA	0.856 *	0.629*
Sb	0.394 *	0.101
SbA	0.210	-0.201
A	-0.141	0.515 *
AConifer	0.219 *	0.081
Analysis set 2		
S,SA	0.278 *	0.072
P, PA	0.675 *	0.126
Sb, SbA	0.282 *	-0.113
A, AConifer	0.091	0.260 *

*significant at 0.20 level

Table 11: Partial Correlations.

	Correlation Coefficients			
	height/age		density/age	
Label Controlling	density		height	
Analysis set 1				
S	0.507	*	0.122	
SA	0.151		0.173	
P	0.447	*	-0.025	
PA	0.802	*	0.438	
Sb	0.480	*	0.193	
SbA	0.141		-0.484	
A	-0.436	*	0.449	*
AConifer	0.030		0.067	
Analysis set 2				
S,SA	0.325	*	0.166	
P, PA	0.654	*	-0.009	
Sb, SbA	0.341	*	-0.115	
A, AConifer	-0.190		0.294	*

*significant at 0.20 level

include the poor accuracy of the forest cover map, and the fact that class data were used rather than continuous data. The extensive fire history and variable growth conditions also contributed to the low correlations obtained. Height and density correlations were also calculated but figures were consistently lower, and therefore not pursued further.

The rejection of subhypotheses A and B results in rejection of the main hypothesis with the conclusion that a forest and land cover legend system could be developed. The proposed methodology to follow was to determine if a predictive regression could associate density with age, and height with age. The lack of consistent, significant correlations however, made it difficult to associate any particular height and age class, and density and age class by species. The simplified legend was therefore based only on species and broad densities (i.e., medium- up to 40%: dense- 40% to 80%), which were also among the main factors influencing spectral reflectance determined by Beaubien (1979).

5. CLUSTER ANALYSIS

The objective was to evaluate a multivariate statistical technique called cluster analysis, as an aid to determine the number and possible content of output clusters (subobjective 4). It was therefore proposed as a technique to assist with unsupervised signature generations, and classifications performed on the ARIES-II. Three phases of study were undertaken including a review of cluster analysis algorithms, the determination of the number of clusters, and their identification by geographic reference with the forest cover map. The latter completes the cluster analysis by converting clusters into informational classes by identifying the appropriate ground cover.

The simulated TM data were in a form that could be subjected to cluster analysis - a data base defined as a rectangular region with r rows (scan lines) and c columns (pixels) of one resolution element (case or pixel) (Duran and Odell, 1974). Each pixel consisted of a p x 1 feature vector with p equal to the number of spectral bands. The data base could also be explained as a Cartesian coordinate system with each pair of coordinates (r,c) representing the location of a single pixel (Titus, 1979). The data base was therefore in convenient form as it was digital, and could be related to geographic coordinate systems (e.g., by affine coordinate transformation).

5.1 Methods

The methods employed began with a geographic referencing procedure to place the data in a form that was suitable for the cluster analysis, and to produce a system to locate ensuing clusters on the forest cover map.

5.1.1 Geographic referencing

The digital image was stored on disk on the ARIES-II and was referenced by a scan line and pixel coordinate system. A FORTRAN computer program (Appendix 3) was written, to display the image data as numeric values for each pixel in a user-specified data file, on the Amdahl 5860 computer under the Michigan Terminal System (MTS) at the University of Alberta. The program was written so that the feature vector could be recreated for all or any systematic number of lines and pixels. Each pixel location was assigned a consecutive case number according to its position in the data file. The feature vector for specified line and pixel intersections was therefore stored by case number, and was the source of data for the cluster analyses. The program output included the case number, line and pixel number, and the feature vector.

Cluster analyses result in clusters whose numbers are groups of pixels from the data file. The UTM coordinates of these pixels and thus their locations could then be determined by counting between UTM grid lines on the displayed image though this would be slow and tedious. A

more rapid and accurate system was therefore developed to geographically locate classified pixels by computing their UTM coordinates. The cluster could then be identified more quickly on the forest cover map. This method would also be useful for the unsupervised clustering to locate the UTM positions of classified pixels. To produce this system, the affine transformation was employed with the same set of control points used to create the geometrically-corrected resampled image. The geometric correction and polynomial transformation routines on the ARIES-II were alternately implemented until a satisfactory list of residuals (i.e., approximately equal deviations from their true values with no outliers) was obtained. The output coordinates from the polynomial transform routine (Appendix 4) were then used to compute the transformation equations using version 9 of the SPSS multiple linear regression subprogram. The calculated coefficients of the transforms were then extracted and combined with COMPUTE and LIST CASES commands in SPSS. This produced a tabular listing of UTM coordinates by case number for the previously created data file. The command files to compute the transforms and the UTM coordinates are given in Appendix 5.

5.1.2 Cluster analysis procedures

Cluster analysis was used for determining groups or clusters using pixels as points in multidimensional space. Clustering assumes no knowledge of probability density

distributions of the feature vectors, and formulates the problem as one of partitioning the patterns of pixels into clusters based on a measure of similarity (Moik, 1980). In this study, the measure of similarity for measuring distance between points was Euclidean distance (Goldberg *et al.*, 1975; Swain and Davis, 1978). The Euclidean distance between two pixels in a total space of n spectral bands is:

$$d_{ij} = [\sum_{t=1}^{n} (x_{it} - x_{jt})^2]^{1/2} \quad \text{where:}$$

x_i, x_j are projections of pixels i and j on spectral band t where ($t = 1,2,...n$). (Green, 1978).

The procedures that were employed were a modification of those presented by Anderberg (1973) and included:

1. sample selection
2. selection of variables (spectral bands)
3. selection of clustering algorithms
4. computer implementation
5. selection of the number of clusters
6. interpretation of results and placement of information labels.

The previously described FORTRAN program was used to select pixels at increments of 25 lines and 30 pixels. A systematic sample was therefore obtained resulting in 198 pixels or 0.2% of the entire study area. An evaluation based on sampling considerations and clustering procedures employed was also performed to assess the implications of

sample selection.

The selection of variables has the greatest influence on the ultimate results of a cluster analysis (Anderberg, 1973). In this study, all 3 spectral bands that were available for analysis were used although these did not cover the actual TM spectral range.

The premise of this study was to use a particular cluster analysis technique for selecting the number of clusters. The problem however, was to first identify which clustering algorithm would best produce forest cover types from the image data, since different classifications will result from different algorithms. For this study, several hierarchical clustering algorithms were evaluated. In the hierarchical approach, the clustering process proceeds sequentially such that at each step, only one pixel or cluster group changes. Hierarchical techniques are either agglomerative or divisive and all start with the calculation of a similarity matrix (Everitt, 1980). Agglomerative starts with every pixel as a single group with successive pooling into a single cluster (Figure 4). Trivedi *et al.*, (1979) used an agglomerative approach for feature selection and classification. Howarth and Wickware (1983) used a divisive approach to classify spectral data using a general computer. Both approaches have therefore been used in remote sensing although hierarchical agglomerative was better known. For this study, four hierarchical agglomerative methods were evaluated which included single linkage (nearest neighbour),

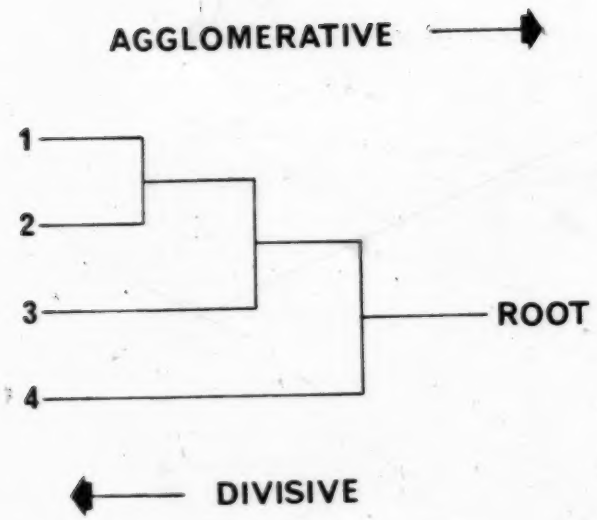

Figure 4: Hierarchical clustering directions for an example with four objects.

average linkage, centroid sorting, and Ward's Method. Single linkage uses the shortest Euclidean distance between two pixels and was suited to finding the closest link between clusters (Volland and Connelly, 1978). This method tends to produce "snake-like" clusters however, even if there were compact clusters in the data set (Anderberg, 1973). The distance measure in average linkage was the average distance from points in the first cluster to points in the second cluster (Green, 1978). In centroid sorting, distances between pixels were weighted by the number of pixels that comprise the cluster (Volland and Connelly, 1978). Clusters were therefore joined according to the shortest centroid distance between them. In Ward's Method, the criterion was to minimize the increase in the total within-group error sum of squares and was evaluated by using the squared Euclidean distance between centroids of the merged clusters (Zupan, 1982). This method shows a tendency to balance out the size of clusters although they could have poor separation properties (Anderberg, 1973).

The statistical package CLUSTAN (Wishart, 1978) on the University of Alberta MTS system was used in this study. The program output lists the members of merging clusters at each step by case number. The case number was then referenced with the data file to identify which pixels were being clustered.

As with unsupervised clustering, a difficulty was determining the most appropriate number of clusters. Formal

significance tests cannot be derived due to the difficulty in specifying a suitable null hypothesis (Green, 1978), and in developing a flexible test procedure (Everitt, 1980). The problem was therefore to determine the stage in the hierarchical clustering process which would be regarded as optimal. This study followed the most frequently used approach by identifying large relative changes in similarity coefficient values between successive merging clusters.

It must be emphasized that cluster results would be a proposition regarding the organization of the data. A solution would be optimal and several solutions could exist. This has important implications in remote sensing application in determining whether or not a set of clusters actually represented forest and land cover types. To perform the evaluation, several pixels in each cluster from the four clustering algorithms were located on the forest cover map using the calculated UTM coordinates. The sets of clusters produced by each algorithm were compared to identify;

a)which algorithm produced clusters that most closely represented land cover types; and

b)the number of clusters needed to classify the area. The methods were summarized in a flowchart (Figure 5) and the results were to provide an information base as a prelude to the unsupervised classification.

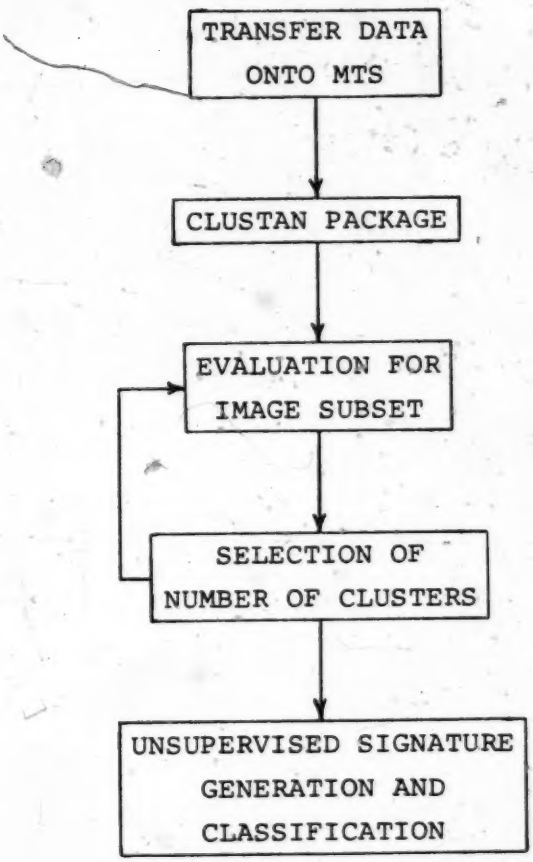
TRANSFER DATA ONTO MTS
CLUSTAN PACKAGE
EVALUATION FOR IMAGE SUBSET
SELECTION OF NUMBER OF CLUSTERS
UNSUPERVISED SIGNATURE GENERATION AND CLASSIFICATION

Figure 5: Flowchart of cluster analysis sequence.

5.2 Results and Discussion

5.2.1 Geographic referencing

Transferring the TM data from the ARIES-II onto MTS resulted in a data file, that could be used for performing the cluster analyses using the CLUSTAN program. A more rapid system was also developed to locate on the forest cover map, classified pixels from the cluster analyses or from the unsupervised clustering using the affine transformation and SPSS. The resulting equations for determining UTM coordinates produced high R^2 values and standard errors of estimates (SE) of about the same magnitudes as for the raw data (Table 12). Control points were however, more difficult to locate on the resampled image. In addition, one control point could not be located leaving a total of nine for generation of the transformation equations. Easting and pixel and northing and line correspond to the same geographic direction. The line and pixel variables for equations 1 and 2, respectively, were not highly significant because the digital image had been geometrically corrected. The equations can therefore be considered scale factors in their relation of the two coordinate systems. The standardized residuals were random with no outliers. This procedure resulted in UTM coordinates of classified pixels which could be located quickly on the forest cover map.

Table 12: Affine transformation results for geographic referencing.

Equation 1: UTM Easting = 444336.06426 + 29.79727 * Pixel + 0.10218 * Line

	Variable	F value	Significance level
R^2 = 0.9997	Pixel	17251.961	0.0000
SE = 68.338	Line	0.386	0.5574

Equation 2: UTM Northing = 6665019.440 - 29.93643 * Line - 0.08018 * Pixel

	Variable	F value	Significance level
R^2 = 0.9998	Line	33991.672	0.0000
SE = 67.448	Pixel	0.128	0.7325

5.2.2 Cluster analyses

In terms of cluster analysis results, the magnitudes and rates of change of the similarity coefficients between the algorithms varied greatly, and were not directly comparable (Table 13). Individual interpretation of each algorithm resulted in selection of 6 clusters for single linkage and average linkage, 7 clusters for centroid sorting, and 8 clusters for Ward's. The number of pixels per cluster differed greatly (Table 14), as did their interpretation with the exception of average linkage and centroid sorting which were similar.

Single linkage resulted in almost all of the pixels other than the 21 in a water class, being in a single cluster (Table 14). Geographic referencing of several pixels in the large cluster to the forest cover map showed no recognizable pattern. The pixels in the large cluster were identified on the map, to be mixed consisting of medium and high density softwoods, treed bog, hardwoods, mixedwoods and

Table 13. Differences in similarity coefficients with the change from relatively small to large values underlined.

	3 Spectral Bands			
# Clusters	Single Linkage	Average Linkage	Centroid	Ward's
11 - 10	.023	.066	.016	.519
10 - 9	.013	.017	.102	.155
9 - 8	.003	.218	.034	.027
8 - 7	.003	.072	.078	1.197
7 - 6	.002	.091	.308	2.253
6 - 5	.083	.305	.085	6.507
5 - 4	.122	.080	.012	4.303
4 - 3	.084	1.865	1.456	25.312
3 - 2	.154	.353	2.055	45.659
2 - 1	.235	2.283	.646	104.118

Table 14: Number of pixels per cluster.

	Cluster number							
	1	2	3	4	5	6	7	8
Single linkage	173	1	21	1	1	1		
Average linkage	10	115	49	1	21	2		
Centroid sorting	11	114	49	1	21	1	1	
Ward's	11	26	19	34	35	50	21	2

grasslands. It was not possible to separate these cover types using single linkage. The chaining effect was therefore not useful for clustering remote sensing data, at least with the data set used.

Average linkage and centroid sorting produced similar results with almost identical pixels in each cluster (Table 14). Other than the same 21 pixels for the water class (cluster number 5), these techniques suggested only a minor increase in definition due to high pixel membership in two clusters with much smaller amounts in the remainder. Further, they were not considered useful in remote sensing since the clusters were not well separated, and did not represent forest land cover types.

The relative magnitudes of the coefficient values for Ward's Method were larger than for the previous methods (Table 15). This facilitated the selection of the number of clusters in comparison to the other methods. There was also a more even distribution of pixels per cluster (Table 16). This confirms allegations by Anderberg (1973) in regard to clustering behavior with Ward's that this method tends to balance out the size of clusters produced. Several pixels from each cluster were located on the forest cover map for identification. Compared to the other three algorithms, Ward's produced clusters which were most representative of mapped cover types. The cover types identified were hardwoods, treed bog, water, several classes of mixedwood, and a general open class. If the water class was deleted,

there would be 6 or 7 land cover classes although separate identification for forest cover was not possible.

Two possible reasons for the difficulty in identification were in sample and variable selection. There were two concerns in sample selection. First, the sample was selected over the entire area which included water and cultural features such as the airport, town of Fort Smith, and roads. The data file therefore contained greater spectral variability than what was of interest. Secondly, pixels were selected systematically at fairly large, specified intervals. This procedure may have inadvertently resulted in varying sizes for the output clusters, which would make them more difficult to analyze (Duran and Odell, 1974). Anderberg (1973) suggested selectivity of data units in order to obtain a sizeable enough sample from small groups for classification. Cluster groups in the data would otherwise tend to be represented in proportion to their relative size in the total sample. This concern may have been alleviated if sampling was performed at smaller intervals resulting in a larger data file, or with blocks of contiguous pixels.

In terms of variable selection, airborne MSS bands 5, 7, and 9 do not cover the full range of spectral information to be covered by the TM. In addition, bands 5 and 7 were highly correlated (0.951). The spectral bands used were therefore a limiting factor in the cluster analysis.

Despite these concerns, interpretation of cluster analyses using Ward's did suggest 6 or 7 land cover classes which would include hardwoods, treed bog, mixedwood, water, and several other classes of land cover. These results formed the basis for the unsupervised classification on the ARIES-II.

6. UNSUPERVISED CLASSIFICATION

Unsupervised classification (signature generation and classification) was rationalized (section 2.3) as the most feasible method for classification since there was little or no prior information on class structure. The two difficulties associated with unsupervised classification were estimating the number of classes, and assigning meaningful land cover labels (Fox and Mayer, 1979). For the former, the initial number of spectral classes was estimated to be twice the number of land cover categories required (Fox and Mayer, 1979; Townshend and Justice, 1980). The number of land cover categories was obtained from the cluster analysis results. The second difficulty was addressed through the geographic referencing procedure developed for the cluster analyses. Classification was followed by filtering.

6.1 Methods

6.1.1 Unsupervised signature generation and classification

Signature generation was performed on the ARIES-II which utilized a migrating means clustering algorithm (Letts, 1978). This parametric algorithm defines clusters using the average Euclidean distance from each pixel to the corresponding pixel mean (Goldberg *et al.*, 1975). This procedure began with the creation of a maxima file, followed

by generation of signature files for use in a three band maximum likelihood (Lillesand and Kiefer, 1979; Kalensky *et al.*, 1981) classification. The bands used were the simulated TM bands. The requirement to specify several parameters (e.g., feature limits, bin sizes, purity filter, and merge factor described below) resulted in considerable latitude to influence a classification. The difficulty was however, in determining the appropriate responses to these parameters. Signature generation and classification were therefore completed several times to compare results, and to become familiar with how the digital data were being clustered (DIPIX Systems Ltd., 1982). The procedure employed included:

1. The generation of histograms from the image data for each spectral band (feature file) to specify feature limits. This was followed by specifying the number of values to allocate to each histogram (bin sizes). Histograms were created for the entire study area, and for a subarea of only vegetated land cover. When specifying values, larger bin sizes produce fewer histogram peaks resulting in fewer themes after classification. A purity filter can also be specified which discards pixels in histogram generation, and influences the pattern of histogram peaks.

2. The generation of the 32 highest peaks or local maxima (uni-modal classes) from the multidimensional histogram, with each peak identified by the number of pixels and their mean and standard deviation in each band.

3. The selection of a merge factor to join the uni-modal classes according to their degree of overlap in the 3 bands, and production of a maxima file. This procedure governed the number of signatures and in turn, the number of themes produced from classification. The selection of the merge factor in determining the number of themes was functionally analogous to the similarity coefficient value in cluster analysis.

4. Production of maximum likelihood signatures according to the classes in the maxima file.

5. Performing a maximum likelihood classification.

6. Displaying the classification and determining the information label for each theme class produced. The latter was performed using the geographic referencing procedure described earlier (section 5.1.1). Identification was difficult since precise location was inhibited by the small size of many land cover units. This problem was further compounded by the nominal two pixel error obtained during the geometric correction. Only those classified pixels that fell within a cover type on the map were therefore used for describing clusters. Any of those that fell near or on the edge of a cover type boundary was not used for description. Generally 5 - 10 pixels of each spectral class were located for identification. These pixels were located by computing their UTM coordinates from the affine transformation equations (section 5.2.1).

7. Performing the next set of unsupervised analyses and comparing them to previous results.

8. Selecting the classification which produced themes most closely representing land cover types in terms of description and performing digital filtering.

A total of 9 analyses were performed (Table 15) to determine which combination of parameters (analysis set) would result in themes most closely representing land cover types.

6.1.2 Post-classification filtering

Classification often results in themes that appear confusing, and which contain more detail (many pixels of different themes in close proximity) than could be interpreted readily (Davis and Peet, 1977). Filtering results in more regular boundaries of themes and provides a closer approximation to what would be produced manually (Davis and Peet, 1977). This procedure is similiar to the cartographic concept of minimum areas used in producing forest cover maps.

Table 15: Unsupervised analyses performed on the ARIES-II

Number	Image set	Histogram limits	Bin size (in each band)	Merge factor	Purity filter used	Signature file created	Theme file created
1	whole image	BNØ9 50-180 BNØ7 40-115 BNØ5 44- 92	4	0.5	no	yes	no
2	whole image	as above	6	0.5	no	yes	yes
3	vegetated area	BNØ9 42-175 BNØ7 40- 76 BNØ5 44- 70	6	0.5	no	yes	yes
4	vegetated area	as above	5	0.5	no	yes	no
5	vegetated area	as above	7	0.5	no	yes	no
6	vegetated area	as above	4	0.5	no	yes	yes
7	vegetated area	as above	4	1.72	no	yes	no
8	vegetated area	as above	4	1.42	yes,3	no'	no
9	vegetated area	as above	4	0.5	no	yes'	yes

'Signature file not created with parameters changed for analysis number 9.

'manual edit of local maxima prior to signature generation

The post-classification minimum area filter on the ARIES-II can be employed on all themes or specified individually for each theme (Letts, 1979). The effects of filtering were determined by comparing themes visually, and by tabular UTM summaries. Filtering consists of two stages including the creation of an area map, and the reassigning of areas less than a specified minimum (DIPIX Systems Ltd., 1982). With the former, each pixel is assigned a value equal to the size of homogeneous area to which it belongs. With the latter, areas less than a specified minimum are replaced equally from all sides by the surrounding pixels. Filtering required the specification of several parameters including the option to include diagonally connected pixels, the maximum replacement depth, and the minimum area for each theme or all themes. Three filtering analyses were performed to identify which combination of parameters, filtered the themes more effectively in terms of physically representing land cover types (Table 16).

6.2 Results and Discussion

6.2.1 Unsupervised signature generation and classification

The parameters used in the first analysis resulted in a maxima file with a minimum of 30 or more pixels in all 32 local maxima classes. The initial objective was to reduce the number of local maxima to approximately 14 classes (2x cluster analysis result). Adjustment in several parameters

Table 16: Post-classification filtering performed on the ARIES-II.

Number	Diagonal pixels included	Replacement depth (pixels)	Minumum area for themes(pixels)
1	yes	7	4 for all themes
2	no	7	4 for all themes
3 and subjectively	yes	7	ranging from 3 - 20 depending on theme selected

would achieve this by producing local maxima with either no or small numbers of pixels. A classification was not performed and bin size was adjusted for the next analysis.

A bin size of 6 in analysis 2 resulted in 10 local maxima having less than 10 pixels. This would reduce the number of local maxima used to produce signatures. Larger merge factors would also have reduced the number of classes. This would inhibit however, observation on which classes would have been represented by the same theme. Signature generation produced 24 signatures which required 3 theme files. Several pixels were located for each theme, although they were very mixed with several classes for water and an airport which was of no interest. The large amount of spectral variability contributed by water and cultural features had to be removed.

Analysis 3 was based on a vegetated subarea from which classification results were more defined for forest land cover types. The range of feature values for each band from the subarea histogram was narrower (Table 15), and indicative of the reduced spectral variability in comparison to the entire study area. With the subarea file and the same bin size, there were also much fewer local maxima classes which resulted in 11 themes. Five of these themes were scattered pixels around Fort Smith and the airport. The remaining themes were of vegetation (mixed softwood, softwood, grassland/open bog, mature spruce, and 2 classes of hardwoods), although they were not complete since some open grassland areas were not classified. The softwood theme was also quite mixed with several densities and species of softwoods. More classes were therefore needed to separate some of these themes.

A bin size of 5 was used in analysis 4 with the maxima file marginally listing more classes. A classification was not performed since the increased classes were small and only contained 1 to 5 pixels. A bin size of 7 was also employed (analysis 5) although results were expected to be fewer classes. Results were confirmed with the maxima file only listing 9 classes. This combination of parameters would result in more mixed themes since analysis 3 with 11 classes already contained mixed themes.

A bin size of 4 was evaluated in analysis 6 which in comparison to earlier analyses, resulted in more themes, and

a larger number of pixels in the smaller themes. The same merge factor was employed to determine if there were more pure themes, and which themes were classifying the same cover types. These parameters resulted in 16 signature files and 2 theme files. This classification was most complete in terms of cover types compared to previous analyses. There were several themes representing the same cover types (Table 17), and these were reassigned the same color (Plate 2, left image). A bin value of 4 and classification based on feature limits from a training area histogram (Table 15), resulted in multiple themes with several of no significance (Table 17). This classification became the working base to evaluate the effects of merge factors (analysis 7, 8), purity filter (analysis 8) and the edit function (analysis 9).

The subsequent analysis (number 7) was to first determine if a large merge factor (1.72) would reduce and join the appropriate classes of the same theme (Table 17). The number of 7 classes was obtained although many classes were merged into 2 classes with the others unchanged. The difficulty observed with merge factors was the analyst not having control to join or delete classes. A classification was not pursued since more mixed themes were anticipated.

The effects of a purity filter and a merge factor to obtain 7 classes (1.42) greatly influenced the pattern and merging of local maxima (analysis 8). A larger purity filter supposedly assists classification of very homogeneous areas (DIPIX Systems Ltd., 1982), although this may be of little

Table 17: Themes produced from unsupervised analysis number 6.

Theme no.	Color	Feature
1	red	dense softwood (40-80% crown closure)
2	green	medium density softwood (to 40%)
3	medium blue	mature softwood (e.g. age to 120 years) mostly black spruce
4	yellow	mixedwood
5	magenta	treed bog and some softwood-hardwood mixed stands
6	cyan	anomaly of image due to haze, class not considered
7	orange	open bog and grassland
8	purple	hardwood
9	orange	same as 7
10	orange	same as 7
11	purple	same as 8
12	magenta	same as 5
13	magenta	same as 5
small classes of no significance to study		
14		airport runway - few pixels
15		vegetation along roads to airport - few pixels
16		few scattered anomalous pixels in open areas

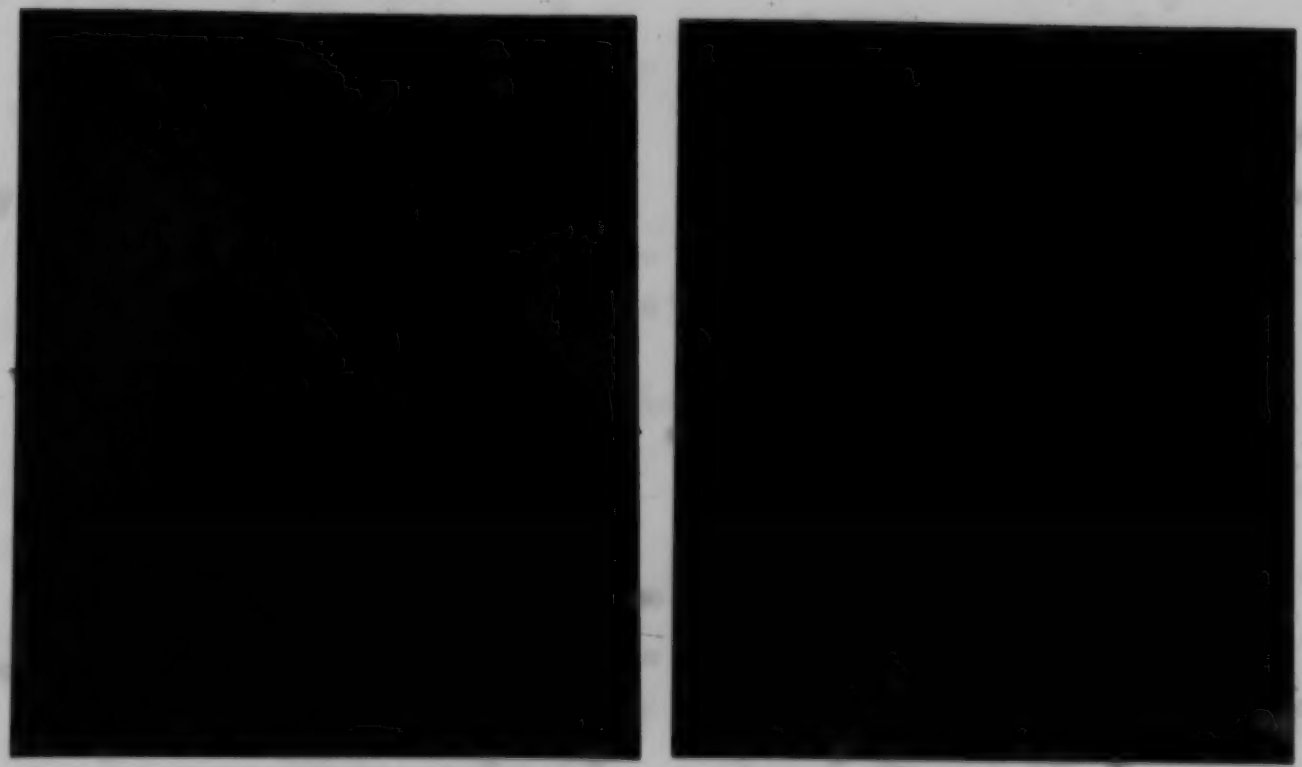

Plate 2: Unfiltered (left) and filtered (right) unsupervised results with multiple themes combined (reference Table 17).

use in complex vegetative areas. Merging of the "new" local maxima resulted in all classes merging into a single class. This approach was anticipated to result in several small themes, and a single large theme consisting of many cover types if a classification was performed.

Analysis 9 employed the edit function to manually merge appropriate classes of the same theme, and to delete unwanted classes (Table 17). The classification visually showed more treed bog and open grassland in comparison to analysis 6 which visually had more mixedwood. In addition, the light blue analomy due to haze on the left side of the study area (Plate 2, left image), was merged with the hardwood class although it was manually merged with a different class. These results confirm previous applications that hand-merging classes has been unreliable (Letts, 1978). In addition, the intended use of the edit facility was more for excluding insignificant classes from the signature generation process (DIPIX Systems Ltd., 1982), than it was for merging selected classes.

In summary, analysis 6 was the most effective for classifying the desired forest land cover types of those analyzed. Attempts in altering parameters to produce single themes were not successful. Results confirmed unsupervised classification behavior, that more than one theme can correspond to the same cover type, or one theme can represent several cover types (Letts, 1978). The lack of control on merging classes using a merge factor suggests

using as small a value as possible. This will keep all classes intact so that observation can be made on which classes represent the same ground cover. The most practical approach is to reassign colors to identical themes, and compute UTM summaries as combinations of the appropriate themes. Deleting unwanted classes in analysis 6 resulted in 7 themes which was equal to the number from cluster analyses.

In comparison, there were 6 classes produced from the previous in-house LANDSAT digital classification project. These classes included spruce softwood forest, pine softwood forest, hardwood forest, shrub/grassland/bog, open black spruce, and open areas with snow. The last theme included cultural features, water, and other land cover classes such as grassland. This classification included confusion in grassland areas, and did not account for mixedwood areas.

6.2.2 Post-classification filtering

The selection of parameters for postclassification filtering can significantly alter the distribution of pixels in each theme. In addition, the degree of filtering considered satisfactory is judgmental.

The first filtering run reduced the complexity of the image and resulted in more pure themes (Plate 2, right image) when compared to the unfiltered and medium density softwood (green) themes in the centre of the images.

The second filtering run was to discern what differences would occur if diagonally connected pixels were not included. Results were more pure themes and although more aesthetically pleasing, the image was suspected to be less realistic. This was due to a large reduction in heterogeneity. These results were an example of how themes could be misrepresented by over smoothing. Filtering should therefore be used with caution since results will influence tabular summaries.

The third filtering run was an attempt to delete and convert to unclassified the unwanted themes, and to vary the minimum area of other themes. The latter was subjective with the values ranging from 3 to 20 pixels, and dependent on some of their small occurrences on the image. The results included a reduction of the unwanted pixels, although the overall appearance was very mottled and speckled. Without guidelines to specify the appropriate minimum area for each theme, it was difficult to assess filtered results. The first analysis was considered the most reasonable of those performed, based on the relative appearances of the filtered results. Post-classification filtering reduced the number of unclassified pixels from 33 823 to 29 435 (from 23.4 to 20.5% of the study area). Mixedwood and softwood cover types were concluded to be the major land cover types, with wetlands also occurring in large amounts (Table 17 and Plate 2). In addition, there appears to be relatively few large stands of pure hardwoods.

7. SUMMARY AND RECOMMENDATIONS

Airborne scanner data were acquired to simulate data from the Thematic Mapper (TM), for forest and land cover classification in the NWT. Several concerns with respect to radiometric, spectral, and spatial resolution were identified in using such data, with the conclusion that they should be as representative of the TM as possible. The radiometric resolution was the same as for the TM, and spectral resolution was simulated by selecting the appropriate wavelength bands. The availability of only 3 of the 6 reflected TM spectral bands was a limitation in this study. TM spatial resolution was simulated by performing a geometric correction followed by pixel resampling. The lack of 1:50 000 NTS maps of the study area and easily recognizable control points (e.g., road intersections and hydrographic features that aren't subjected to seasonal changes), negated the employment of frequently used geometric correction procedures. Identification and transfer of control points, using the small scale (1:66 000) photos and radial line triangulation, was determined as the only approach by which geometric correction could be completed. Even if 1:50 000 NTS maps had been available, the procedures employed would still have been necessary due to the difficulty in locating control points.

Correlation coefficients between height and age, and age and density, were generally low, and were not consistently significant for all species groupings. This

made it difficult to relate spectral image data to forest cover information, and thereby develop a legend to describe clusters from classifications. The data were also not sufficiently continuous in all species groups for this approach to be implemented. The legend was therefore based on the two main factors of species groupings and density.

The unsupervised approach to classification was selected because there was little prior information on class structure. A difficulty with this approach however, was in determining the final number of meaningful output clusters. A multivariate cluster analysis technique was therefore evaluated as an aid in determining the number, and possible content of output clusters when performing unsupervised clustering on the ARIES-II. A cluster analysis statistical package called CLUSTAN was available on the MTS operating system at the University of Alberta. A program was therefore developed to create MTS data files from image files stored on tape. A geographic referencing technique, based on an affine transformation on the geometrically-corrected image, was also developed to determine the UTM coordinates of classified pixels from cluster analyses. This technique was useful in identifying output clusters, although it was a slow and tedious procedure to physically locate these clusters on the forest cover map. Of the clustering algorithms evaluated, Ward's method produced clusters which most closely represented land cover types. Interpretation resulted in 7 clusters (hardwoods, treed bog, mixedwood,

water and several classes of land cover), although the exact composition of all clusters, was impaired by using sample data over the entire study area (including water and the airport).

Unsupervised signature generation and classification on the ARIES-II required the specification of several parameters including feature limits, bin sizes, purity filter, and merge factors in generating signatures. This resulted in considerable latitude to influence a computer classification. The unsupervised process must be performed several times to compare results, although a guide of approximately 2x the final number of land cover classes (from cluster analysis), was a reasonable number to start with. It is suggested to use as small a merge factor as possible to keep classes intact, and to restrict use of the edit function to only deleting unwanted classes. Pixels that were classified were located on the forest cover map with the geographic referencing technique used in the cluster analyses. A vegetated subarea was created to reduce the spectral variability from water and cultural features, which resulted in more defined classes than when the entire study area was considered. The occurrence of multiple and mixed themes confirmed earlier studies (Letts, 1978), and were concluded to be indicative of classification behaviour with unsupervised approaches. This was especially prevalent in this study due to the inherent complexity of vegetation with respect to species, density and other variables such as age,

height, and other ground vegetation. The final 7 themes were determined to be dense softwood (40-80% crown closure), medium density softwood (to 40%), mature softwoods of mostly black spruce (up to 120 years), mixedwoods, treed bogs and some softwood-hardwood mixed stands, open bogs and grasslands, and hardwoods. Minimum area filtering was a necessary exercise to produce contiguous areas of themes, rather than to have a classification that appeared very mottled and speckled. Caution must be exercised however, to prevent misrepresentation by over smoothing. A tabular summary of themes for the entire study area was produced, although breakdowns by UTM cell is also available on the ARIES-II.

The use of simulated Thematic Mapper data resulted in species groupings and broad density classes to describe themes, whereas the previous in-house project with LANDSAT MSS data only used species. This investigation into potential gains in using higher resolution data however, was limited by the low reliability of the reference information available, and the few airborne spectral channels that were used. The level of refinement was therefore conservative in relation to what has been reported in the literature. The location and vegetative complexity of the study area, required the use and development of methods not previously employed in such a combination for forest and land cover classification.

7.1 Recommendations

Several concerns arose during the course of this research. Further investigation is therefore suggested to address several other areas of research.

1. What would be the total quantitative effects on classification summaries by resampling pixel size to 25 m instead of 30 m? This pixel size would result in an even number of pixels between UTM grid lines, and be similiar in concept to resampling LANDSAT MSS data from 79 m to 50 m.
2. For the geometric correction process, how much reduction in location errors could be expected from the use of 1:50 000 NTS maps, and the methods described in this study in northern boreal environments? The procedures developed are feasible for operational use although accuracies are marred until larger scale maps become available.
3. The correlation analysis approach in producing a simplified legend system for describing clusters, should be further evaluated in an area where forest cover has been accurately documented.
4. Would an alternate sampling strategy of selecting pixels for cluster analyses, improve the definition of the number and identification of output clusters? The suggested approach is to perform a separate cluster analysis on individual contiguous blocks of pixels. Statistics of selected clusters could then be used to

classify the entire study area. The advantage would be in focusing attention on forested areas where maximum discrimination is desired.

5. Would a refinement to the selection of vegetated areas for unsupervised classification result in more refined classes? The approach employed in this project improved definition of classes in comparison to classifying the entire study area. An extension of this approach would be the multi-cluster blocks technique (Hoffer and Fleming, 1978; Nelson and Hoffer, 1979), whereby unsupervised clustering is performed on several selected areas. Selected signatures generated from these areas are then used to classify the whole area. This suggestion is similiar in concept to number 5.
6. The location of control points and classified pixels on reference maps was a slow and tedious procedure. Although techniques were developed to assist these tasks, the most efficient approach would be to incorporate a digitizing work station, with appropriate software, into the ARIES-II system.

7.2 Proposed Role of Satellite Remote Sensing in the NWT

The suggested role of satellite remote sensing is not to replace but supplement existing methods used in completing operational surveys. The projected scenario is the production of a contiguous information base which will guide where more detailed forest inventory operations should

take place, and which would also be used to assist fire management decisions. For example, fire location and direction data from aerial survey spotting and mapping, could be transferred onto the forest cover information base to assist determination of projected fire damage. It is therefore hoped that planning and fire fighting deployment decisions could be aided using the described combination of data. It is also projected that once a complete forest inventory data base has been produced, that the major role would revert to monitoring of the forest resource for alterations due to for example, by fire and harvesting activity.

The incorporation of relatively new techniques often requires several demonstrations, and several years of development work before they can be accepted (Harding, 1982). This study was therefore, not expected to result in an immediate operational technology transfer. Rather, it is presented for the DIAND to detemine if such remote sensing techniques merit consideration for use in their forest inventory programs.

BIBLIOGRAPHY

Ahern, F.J., Brown, R.J., Goodenough, D.G. and Thomson, K.P.B. 1980. A simulation of thematic mapper performance in an agricultural application. pp. 585 - 595 in T.T. Alfoldi (ed.), Proc, 6th Canadian Symposium on Remote Sensing, Halifax, Nova Scotia.

Airborne Operations Section. 1977. Information bulletin:airborne operation. Energy, Mines and Resources Canada, Canada Centre for Remote Sensing, Data Acquisition Division. Ottawa, Ont. 3rd edition.

Aldrich, R.C. 1979. Remote sensing of wildland resources: a state-of-the-art review. USDA, For. Serv., Rocky Mtn. For. and Rge. Expt. Sta., Fort Collins, Co. Gen. Tech. Rep. RM-71.

Anderberg, M.R. 1973. Cluster analysis for applications. Academic Press, Inc., New York, N.Y.

Audet, H. and Thomson, K.P.B. (eds.). 1981. Colloquium on the application of the next generation of earth resource satellites. Gouvernement du Quebec, Ministere de l'Energie et des Ressources, Montreal, Quebec.

Beaubien, J. 1979. Forest type mapping from LANDSAT digital data. Photogramm. Eng. Remote Sensing 45(8):1135-1144.

Bickerstaff, A., Wallace W.L. and Evert F. 1981. Growth of forests in Canada Part 2: A quantitative description of the land base and the mean annual increment. Env. Can., Can. For. Serv., Petawawa Nat'l For. Inst., Chalk River, Ont. Inf. Rep. PI-X-1. 136pp.

Bonnor G. M. 1982. Canada's forest inventory 1981. Environ. Can., Can. For. Serv., For. Stat. and Systems Bran. Chalk River, Ont. 79pp.

Bowen, M.G. 1978. Canada's forest inventory - 1976. Environ. Can., Can. For. Serv., For. Manage. Inst. Ottawa, Ont. Inf. Rep. FMR-X-116. 63pp.

Colwell, J.E. 1974. Vegetation canopy reflectance. Remote Sensing of Environment 3:175-183.

Colwell, R.N. 1978. Remote sensing as an aid to the inventory and management of natural resources. Cdn. Surveyor, 32(2):183-203.

Covault, C. 1982. LANDSAT-D to yield more precise data.

Aviation Week and Space Technology 117(1):40-51.

Crist, E.P. 1984. Comparison of coincident LANDSAT-4 MSS and TM data over an agricultural region. pp. 508-517 in Proc. 50th Annual Meeting-American Society of Photogrammetry, Washington, D.C.

Crown P. H. and Pawluk, S. 1973. Spectral reflectance from selected targets at the Kananaskis forest experiment station. Env. Can., Can. For. Serv., North. For. Res., Cent., Contract Rep. Serial Edmonton, Alta. No. OSP3-o223. AIP G-73-1.

Davis, W. A. and Kenue, S. K. 1977. On the detection of changes in digital images. University of Alberta, Dept. of Computing Science, Edmonton, Alberta. Tech. Rep.

Davis, W. A. and Peet, P. G. 1977. Recognition and modification of areas less than a minimum. pp. 549-557 in K.P.B. Thomson (chairman). Proc. 4th Canadian Symposium on Remote Sensing, Quebec City, Quebec.

Dean, M. E. and Hoffer, R. M. 1982. An evaluation of Thematic Mapper simulator data for mapping forest cover. pp. 300-307 in D. C. McDonald and D. B. Morrison (eds). Proc. 8th Machine Processing of Remotely Sensed Data Symposium. Purdue Univ., LARS, West Lafayette, Ind.

Derenyi, E. and Yazdani, R. 1981. Potentials of LANDSAT-D and SPOT-1 for crop identification in the maritimes. pp. 120-125 in W. G. Best and S. Weselake (eds.), Proc. 7th Canadian Symposium on Remote Sensing, Winnipeg, Manitoba.

DIPIX Systems Ltd., 1982. ARIES-II application software package - A2ASP version operator's manual. Ottawa, Ont.

Duran, B. S. and Odell, P. L. 1974. Cluster analysis:a survey, Lecture notes in economics and mathematical systems, Springer-Verlag, New York, New York.

Everitt, B. 1980. Cluster Analysis, 2nd ed., Halsted Press, New York, N.Y.

Forshaw, M. R. B., Haskell, A., Miller, P. F., Stanley, D. J. and Townshend, J. R. G. 1983. Spatial resolution of remotely sensed imagery:a review paper. Int. J. Remote Sensing 4(3):497-520.

Fox, L. and Mayer, K. E. 1979. Using guided clustering techniques to analyze LANDSAT data for mapping forest land cover in Northern California. pp. 364-367 in I.M. Tendam and D.B. Morrison (ed.), Proc, 5th Machine processing of Remotely Sensed Data Symposium, Purdue

Univ., LARS, West Lafayette, Ind.

Freden, S. C. and Gordon, F. 1983. LANDSAT satellites. pp. 517-570. in R.N. Colwell (editor) manual of Remote Sensing. American Society of Photogrammetry, Falls Church, Virginia.

Goldberg, M., Goodenough, D. and Shlien, S. 1975. Classification methods and error estimation for multispectral scanner data. pp. 125-131 in G. E. Thompson (ed.), Proc. 3rd Canadian Symposium on Remote Sensing, Edmonton, Alberta.

Goodenough, D. G., Narendra, P. M. and O'Neill, K. 1978. Feature subset selection in remote sensing. Cdn. J. Remote Sensing 4(2):143-148.

Green, P. E., 1978. Analyzing multivariate data. The Dryden Press, Hinsdale, Ill.

Hall R. J. 1978. A comparison between panchromatic and color vertical aerial photographs in planimetric map construction using the radial planimetric (Kail) plotter. Univ. of B. C., Fac. of Forestry, Unpubl. B. Sc. F. thesis. 79 pp.

Harding, R. A. 1982. Remote sensing as a tool for forest management. pp. 448-453 in C. J. Johannsen and J. L. Sanders (eds.). Remote Sensing for Resource Management. Soil Conservation Society of America, Ankena, Iowa. Chp. 39.

Haymond, W. D. 1982. Radiometric and geometric correction of high resolution data from space. pp. 137-148 in B. S. Mathur (chairman), Proc. 2nd Nat'l. Workshop on Engineering Applications of Remote Sensing, EMR, Canada Centre for Remote Sensing, Edmonton, Alta.

Hegyi, F. and Quenet, R. V. 1981. Forest classification system considerations pp. 69-74 in T. B. Brann, L. O. House and H. G. Lund (eds.), Proc. In-Place Resource Inventories:Principles and Practises Workshop. SAF, Univ. of Maine, Orono, Maine.

Helms, J. R. and Shain, W. A. 1981. Problems related to the use of remote sensing for inventory and mapping of lower coastal plain forests. pp. 540-542 in P. G. Burroff and D. B. Morrison (eds.). Proc. 7th Machine Processing of Remotely Sensed Data Symposium. Purdue Univ., LARS, West Lafayette, Ind.

Hilborn, W. H. 1981. Evaluation of forest inventory classification in airphoto interpretation. pp. 170-175 in T. B. Brann, L. O. House and H. G. Lund (eds.), Proc.

In-Place Resource Inventories:Principles and Practises Workshop. SAF, Univ. of Maine, Orono, Maine.

Hoffer, R. M. and Fleming, M. D. 1978. Mapping vegetative cover by computer-aided analysis of satellite data. pp. 227-236 in H. G. Lund, J. L. Vernon, P. F. Folliott and D. Robinson (techn. coordinators) Proc. Integrated Inventories of Renewable Natural Resources Workshop. USDA, For. Serv., Rocky Mtn. For. and Rge. Expt. Sta. Fort Collins, Co. Gen. Tech. Rep. RM-55.

Hoffer, R. M. and Swain, P. H. 1980. Computer processing of satellite data for assessing agricultural, forest and rangeland resources. pp. 437-446 in Proc. 14th Congress of the International Society of Photogrammetry - Commission VII, Hamburg, Germany.

Horvath, R., Braithwaite, J. G. and Polcyn, F. C. 1970. Effects of atmospheric path on airborne multispectral sensors. Remote Sensing of Environment 1:203-215.

Howarth, P. J. and Wickware, G. M. 1983. Multivariate analysis techniques for land cover classification using multispectral scanner data. pp. 463-473 in K. P. B. Thomson (eds.), Proc. 8th Canadian Symposium on Remote Sensing, Montreal, Quebec.

Hyde, R. F., Vesper, N. J. and Goldblatt, I. A. 1981. The utility of smaller LANDSAT Ground resolution levels. pp. 475-484 in Proc. 9th Annual Conference:Matching Remote Sensing Technologies and their applications. University of Reading, Reading, England.

Hyde, R. G., Spencer, M. K. and Vesper, N. J. 1983. An estimate of cost-effectiveness of machine processing increased spatial resolution MSS data. Butler Univ., Holcomb Research Institute, Indianapolis, Indiana.

Indian and Northern Affairs Canada. 1983. Canada's north: the reference manual. Ottawa, Ontario.

Kachhwaha, T. S. 1983. Spectral signatures obtained from LANDSAT digital data for forest vegetation and land-use mapping in India. Photogramm. Eng. Remote Sensing 49(5):685-689.

Kalensky, Z. and Scherk, L. R. 1975. Accuracy of forest mapping from LANDSAT computer compatible tapes. pp. 1159-1167 in Proc. 10th International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan.

Kalensky, Z. D., Moore, W. C., Campbell, G. A., Wilson, D. A. and Scott, A. J. 1981. Summary forest resource data from LANDSAT images. Environ. Can., Can. For. Serv.,

Pet. Nat'l., For. Inst., Chalk River, Ont. Inf. Rep. P1-X-5.

Kieth, V. and Cooper, M. 1974. Non-parametric design and analysis, Univ. of Ottawa Press.

Kindelan, M., Moreno, V. and Valverde, A. 1981. Geometric correction of airborne multispectral scanner images. pp. 1539-1549 in Proc. 15th International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan.

Kirby, M. and Steiner, D. 1978. The appropriateness of the affine transformation in the solution of the geometric base problem in LANDSAT data. Cdn. J. Remote Sensing 4(1):32-43.

Krebs, P. V. 1982. Multiresource inventory and mapping of Alaska's wildlands: A cost-effective application of remote sensing. pp. 81-90 in C. J. Johannsen and J. L. Sanders (eds.), Remote Sensing for Resource Management, Soil Conservation Society of America, Ankeny, Iowa.

Landgrebe, D. A. 1973. Machine processing of remotely acquired data. Purdue Univ., LARS, West Lafayette, Ind. LARS Inf. Note 031573.

Latty, R. S. and Hoffer, R. M. 1981. Waveband evaluation of proposed Thematic Mapper in forest cover classification. Purdue Univ., LARS, West Lafayette, Indiana. Tech. Rep. 041581.

Leckie D. G. 1980. Use of polynomial transformations for registration of airborne digital line scan images. pp. 635-641 in T. T. Alfoldi (ed.), Proc. 6th Canadian Symposium on Remote Sensing, Halifax, Nova Scotia.

Lee, Y. J. 1980. Application of temporal LANDSAT digital data for updating the Yukon RRAMS data base using ARIES. pp. 73-80 in T.T. Alfoldi (ed.), Proc. 6th Canadian Symposium on Remote Sensing, Halifax, Nova Scotia.

Lee, Y. J. 1981. Comparison of LANDSAT MSS with simulated LANDSAT-D, SPOT, and airborne MSS data for identification of tree species and land classes. pp. A17-A19 in H. Audet and K. P. B. Thomson (eds.), Proc. Colloquium on the Application of the Next Generation of Earth Resource Satellites. Gov't. of Quebec, Montreal, Quebec.

Letts P. J. 1978. Unsupervised classification in the ARIES image analysis system. pp. 61-69 in Y.J. Lee (chairman), Proc. 5th Canadian Symposium on Remote Sensing. Victoria, B.C..

Letts P. J. 1979. Small area replacement in digital thematic maps. pp. 431 in I.M. Tendam and D.B. Morrison (eds.), Proc. 5th Machine Processing of Remotely Sensed Data Symposium. Purdue Univ., LARS, West Lafayette, Ind..

Lillesand, T. M. and Kiefer, R. W. 1979. Remote sensing and image interpretation. John Wiley and Sons, New York, N.Y.

Long, O. and Hopkins, P. F. 1984. Multitemporal analysis of digital LANDSAT MSS data for forest cover discrimination in central New York State. pp. 543-552. in Proc. 50th Annual Meeting - American Society of Photogrammetry, Washington, D. C.

Mayer, K. E. and Fox, L. 1981. Identification of conifer species groupings from LANDSAT digital classifications. Photogramm. Eng. Remote Sensing 48(11):1607-1614.

McQuillan, A. K. 1975. Benefits of remote sensing in Canadian northern resource development. EMR, Can. Cent. Remote Sensing, Res. Rep. 75-6AX. 76 pp.

McQuillan, A. K. 1978. Applications and potential benefits of LANDSAT-D. Energy, Mines and Resources Canada, Canada Centre for Remote Sensing, Ottawa, Ont. Res. Rep. (Draft). 363 pp.

Merchant, J. W. 1981. Utilization of spatial complexity in computer classified LANDSAT MSS data for multi-factoral thematic mapping. pp. 905-914 in Proc. 15th International Symposium on Remote Sensing of Environment. Ann Arbor, Michigan.

Moik, J. G. 1980. Digital processing of remotely sensed images. NASA, Scientific and Technical Information Branch, Washington, D. C., SP-431.

Mroczynski, R. P., Hoffer, R. M., and Nelson, R. F. 1980. computer-aided processing of LANDSAT MSS data for classification of forestlands. Purdue Univ., LARS, West Lafayette, Ind. Tech. Rep. 101880.

Murphy, J. 1979. Format specifications for Canadian LANDSAT MSS system corrected computer compatible tape. Energy, Mines and Resources Canada, Can. Cent. Remote Sensing, Ottawa, Ont. Res. Rep. 79-2.

Nelson, R. F. and Hoffer, R. M. 1979. Computer-aided processing of LANDSAT MSS data for classification of forest lands. Purdue Univ., LARS, West Lafayette, Ind., Tech. Rep. 102679.

Nelson, R. F., Latty, R. S. and Mott G. 1984. Classifying

northern forests using Thematic Mapper simulator data. Photogramm. Eng. Remote Sensing 50(5):607-617.

Nie, J. H., Hull, C. J., Jenkins, J. G., Steinbrenner, K. and Bent, D. J. 1975. SPSS, 2nd edition, McGraw-Hill Book Co., New York, N.Y.

Orte, F. 1981. Optimal distribution of control points to minimize Landsat image registration errors. Photogramm. Eng. Remote Sensing 47(1):101-110.

Pala, S. and Jano, A. 1981. LANDSAT-based forest mapping in Ontario north of latitude 52° north. pp. 228-232. in W.G. Best and S. Weselake (eds.), Proc. 7th Canadian symposium on Remote Sensing, Winnipeg, Manitoba.

Purvee, R. and Braun, K. 1978. An application of the ARIES system to ground vegetation mapping for forestry. pp. 79-85 in Y. J. Lee (chairman), Proc. 5th Canadian Symposium on Remote Sensing, Victoria, B. C.

Rowe, J. S. 1972. Forest regions of Canada. Environ. Can., Can. For. Serv., Pub. 1300.

Schreier, H., Goodfellow, L.C. and Lavkulich, L.M. 1982. The use of multi-date LANDSAT imagery in terrain classification. Photogramm. Eng. Remote Sensing 43(1):111-119.

Shlien, S. 1979. Geometric correction, registration and resampling of LANDSAT imagery. Cdn. J. Remote Sensing 5(1):74-89.

Silva, L. F. 1978. Radiation and instrumentation in remote sensing. pp. 21-135 in P. J. Swain and S. M. Davis (eds.), Remote Sensing:The Quantitative Approach. McGraw-Hill Inc., USA.

Smith, D. M. 1982. Comparisons of LANDSAT-4 with earlier LANDSATS. pp. 503-512 in J. Foreman (ed.), Proc. ISPRS Commission IV Symposium: Environmental Assessment and Resource Management. ASP and American Congress on Surveying and Mapping, Crystal City, Virginia.

Swain, P. H. and Davis, S. M. 1978. Remote sensing: the quantitative approach. McGraw-Hill, Inc., New York, N.Y.

Teillet, P. M., Guindon, B. and Goodenough, D. G. 1981. Forest classification using simulated LANDSAT-D Thematic Mapper data. Can. J. Remote Sensing 7(1):51-60.

Thiruvengadachari, S. 1981. Cost effective computer processing of airborne scanner data for regional level mapping. Photogramm. Eng. Remote Sensing

47(12):1705-1708.

Titus, S. J. 1979. Design of multivariate surveys for sampling wildland ecosystems, PHD dissertation, Univ. of California, Berkeley, Calif.

Tomlins, G. F. 1981. A comparison of LANDSAT Multispectral Scanner (MSS) data and simulated LANDSAT-D Thematic Mapper (TM) data for wetland monitoring in the Fraser River Estuary, British Columbia. pp. A21-A24 in H. Audet and K. P. B. Thomson (eds.), Proc. Colloquium on the Application of the Next Generation of Earth Resource Satellites, Montreal, Quebec.

Townshend, J. R. G. and Justice, C. O. 1980. Unsupervised classification of MSS LANDSAT data for mapping spatially complex vegetation. Int. J. Remote Sensing 1(2):105-120.

Townshend, J. and Justice, C. 1981. Information extraction from remotely sensed data:a user view. Int. J. Remote Sensing 2(4):313-329.

Trivedi, M., Wyatt, C. L. and Anderson, D. R. 1979. An inter-class feature selection procedure. pp. 358 in D. B. Morrison (ed.), 5th Machine Processing of Remotely Sensed Data Symposium. Purdue Univ., LARS, West Lafayette, Ind.

Tucker, C. J. 1978. A comparison of satellite sensor bands for vegetation monitoring. Photogramm. Eng. Remote Sensing 44(11): 1369-1380.

Tucker, C. J. 1980. Radiometric resolution for monitoring vegetation:how many bits are needed? Int. J. Remote Sensing 1(3):241-254.

Ueno, S. 1982. Effects of resampling schemes on the ground albedo mapping of LANDSAT MSS data. pp. 525-535 in J. Foreman (ed.), Proc. ISPRS Commission IV Symposium:Environmental Assessment and Resource Management. ASP and American Congress on Surveying and Mapping, Crystal City, Virginia.

U. S. Geological Survey. 1979. LANDSAT data users handbook. U.S. Dept. of the Interior, EROS Data Center, Sioux Falls, South Dakota.

U. S. Geological Survey. 1982. LANDSAT data users notes. NASA, EROS Data Centre, Sioux Falls, S. Dak. Issue no. 23.

Volland, L. A. and Connelly, M. 1978. Computer analysis of ecological data, USDA, For. Serv., Pac. NW Region, R6-ECOL-79-003.

Williams, D. L. and Stauffer, M. L. 1980. What can the forestry community expect from LANDSAT-D Thematic Mapper data? pp. 475-492 in Proc. Remote Sensing for Natural Resources. Univ. of Idaho, Moscow, Idaho.

Wishart D. 1978. Clustan user manual. Inter-University/Research Councils Series, Program Library Unit, Edinburgh University, Report 47, 175 pp.

Zupan, J. 1982. Clustering of large data sets. John Wiley and Sons, New York, N.Y.

Zwick, H. H., McColl, W. D. and Edel, H. R. 1980. The CCRS DS1260 multispectral scanner (MSS) pp. 643-648 in T. T. Alfoldi (ed.), Proc. 6th Canadian Symposium on Remote Sensing, Halifax, Nova Scotia.

APPENDICES

APPENDIX 1

Tape format files for 11 channel

CCRS JSC airborne MSS tape

Format files were created to extract suitable image files from the 11 channel tape. The order of spectral bands on tape is given in Figure 6. The bands read with each format file were:

Format file no.	Spectral bands
1	1,4,8
2	2,5,9
3	3,6,10
4	X,7,11

Figure 7 illustrates how the data were arranged in the various records on the tape. The ARIES-II dialogue with responses to create the format files follow (Table 18). Image data were transferred to disk using the previously defined format files with the tape header information not displayed. Four area names were created and were subsequently modified to the same area name.

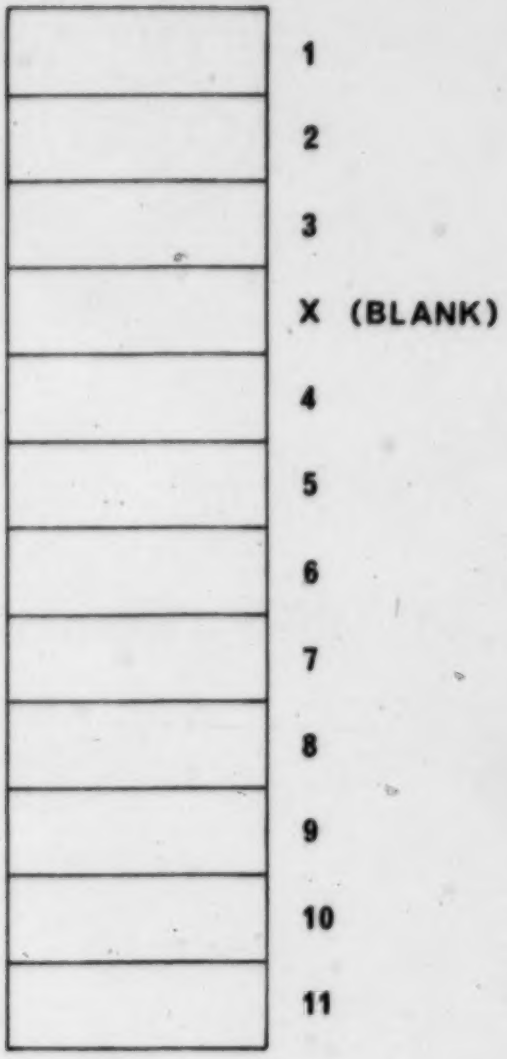

Figure 6: Order of spectral bands on tape.

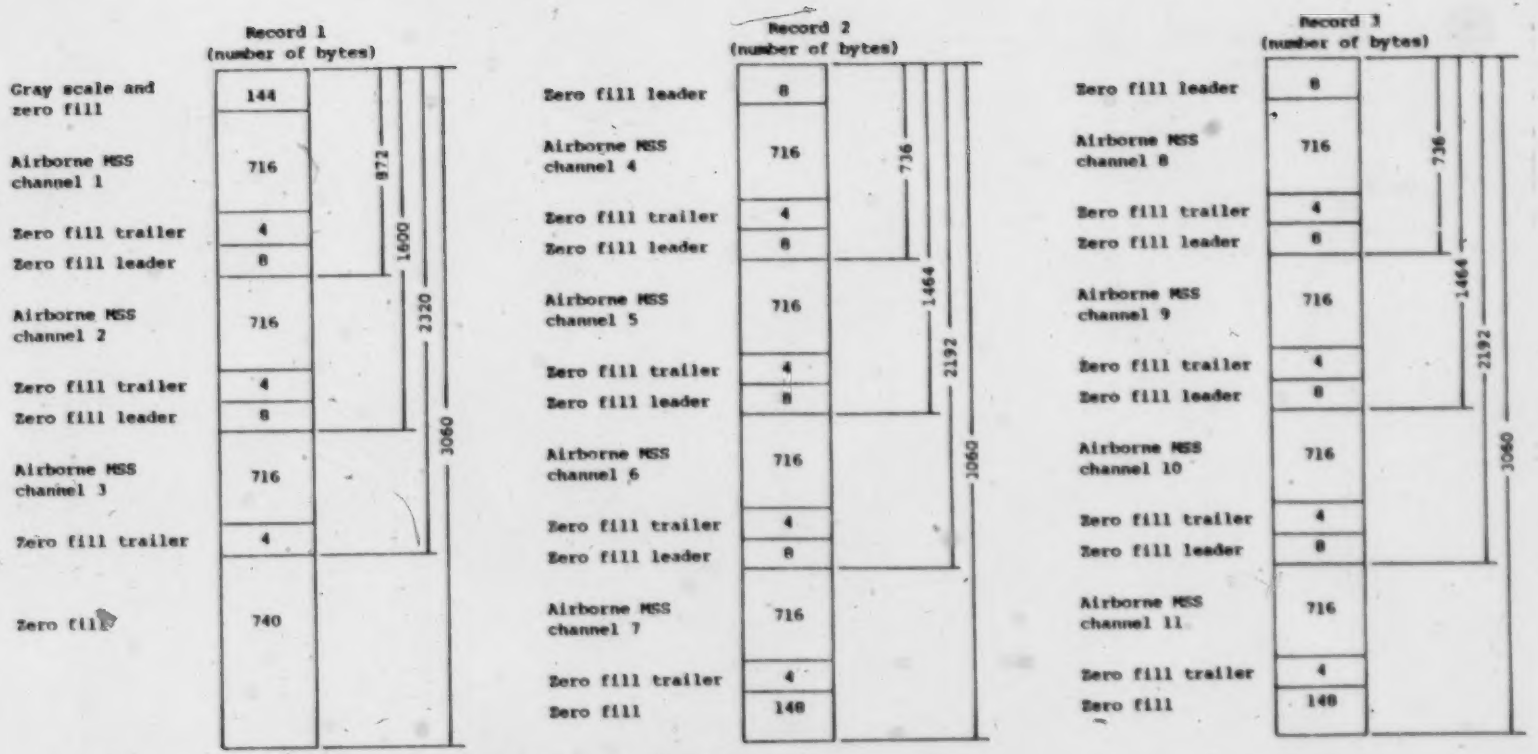

Figure 7. Image data contained in each record on 11 channel tape.

Source: R. Tomkins, DIPIX Systems Ltd., correspondence to K. Campbell, Alberta Remote Sensing Center, Dec. 1982.

Table 18. ARIES-II task TF responses to create format files.

ARIES II dialogue	Format file number			
	1	2	3	4
Name of the format file to be created	AIRMSS001.FMT	AIRMSS002.FMT	AIRMSS003.FMT	AIRMSS004.FMT
All header files on tape will be skipped. Enter number of header files on tape	1	1	1	1
Enter number of features (bands) on tape [1/4]	3	3	3	3
Enter maximum size of physical record on tape in bytes	3060	3060	3060	3060
Enter band number (1 to N) of next band on tape	1	2	3	any number > 11
Enter offset in bytes to first valid data byte in this band, relative to start of physical record (0 or greater)	144	872	1600	2320
Enter number of valid data bytes	716	716	716	716
Enter band number (1 to N) of next band on tape	4	5	6	7
Enter offset in bytes to first valid data byte in this band, relative to start of physical record (0 or greater)	872	736	1464	2192
Enter number of valid data bytes	716	716	716	716
Enter band number (1 to N) of next band on tape	8	9	10	11
Enter offset in bytes to first valid data byte in this band, relative to start of physical record (0 or greater)	1600	736	1464	2192
Enter number of valid data bytes	716	716	716	716

APPENDIX 2

Control points and residual values from
the affine transformation on the uncorrected image

UTM Coordinates

Image Coordinate		Map Easting (m)			Map Northing (m)		
Pixel	Line	Actual	Estimated	Residual	Actual	Estimated	Residual
224	1300	445325.00	445393.09	-68.09	6652987.00	6652996.50	-9.50
316	1173	447238.00	447289.38	-51.38	6655225.00	6655235.50	-10.50
352	1351	448237.00	448171.78	65.22	6652212.00	6652147.00	65.00
557	1373	452650.00	452585.22	64.78	6651763.00	6651826.00	-63.00
612	1058	453500.00	453571.44	-71.44	6657275.00	6657327.50	-52.50
479	1159	450725.00	450779.16	-54.16	6655600.00	6655528.50	71.50
195	933	444575.00	444544.37	30.63	6659300.00	6659378.00	-78.00
519	979	451600.00	451526.69	73.31	6658725.00	6658675.00	50.00
231	729	445225.00	445191.25	33.75	6662975.00	6662941.00	34.00
429	659	449375.00	449397.72	-22.72	6664213.00	6664219.50	-6.50

Standard error of the pixel estimate 62.978m
Standard error of the line estimate 57.058m

APPENDIX 3

FORTRAN program to create data files
from ARIES-II feature files

Purpose of program: To create MTS data files on disk at the University of Alberta from DIPIX ARIES-II image feature files on tape.

Instructions

1. Back off image files from ARIES-II disk onto tape using the back-up and restore routine. The order of files recorded on tape will be the same as the image files transferred from tape onto MTS. Files will be recorded at 1600 BPI and will be 9 track, ASCII. Remember to use the no write ring. The total tape length required to record files is a function of image file size and the number of files to be analyzed.
2. Bring the tape to Computing Science, University of Alberta and register the tape. It is recommended to use the keep ring out option to prevent accidental write on tape.
3. Mount the tape with the user signon ID using the tape rack number assigned.

   e.g. MOUNT 004404 9TP *T*
4. Once mounted, create empty files to record image files.
5. Produce a compiled program called TEST if running for the first time. Assume the FORTRAN program is called TEST.FOR

   RUN *FORTH SCARDS=TEST.FOR SPUNCH=TEST
6. Transfer files from tape onto MTS using the command:

   COPY *T* Filename

Repeating the above will in turn, increment one file on tape. Remember that the first file is a header file. Once copied, the header file may be deleted as it is not needed.

7. Run the compiled program and follow the prompts. The number of lines and pixels and their increment factor, and input and output filenames will be requested. The program has been designed for a systematic extraction of image data according to the line and pixel skip factor used. If all lines and pixels for a certain area is desired, the increment factor will be one. The compiled program will create an output file according to the name specified if it does not already exist. Run the program using the command:

RUN TEST /

8. Before using an output file in any statistical analysis, delete the first line of titles.
9. The output format field is of a FORTRAN type of specification:

1X, NF4.0

where N= line,pixel plus number of image files extracted.

10. The following FORTRAN program is TEST.FOR

```
      SUBROUTINE TEST
      LOGICAL*1 LINE(8192,11)
C     LOGICAL*1 LINE(512,11)
      LOGICAL*1 FILE(40)
      INTEGER*2 LEN
      INTEGER*4 IVAL4(11)
      LOGICAL*1 IVAL1(4,11)
      EQUIVALENCE (IVAL4(1),IVAL1(1,1))
      NUMBND=INUMB('NUMBER OF BAND WIDTHS TO ANALYZE;1-11;')
      LUN=7
      DO 100 I=1,NUMBND
      LUN=LUN+1
      WRITE(6,30)I
   30 FORMAT('& INPUT FILE NUMBER',I2)
      CALL ASSIGN(LUN,'ENTER IN NAME : ;',FILE)
  100 CONTINUE
C **  RATE AT WHICH TO DO LINES AND PIXELS
      ISTL=INUMB('START AT WHICH LINE;')
      INCL=INUMB('INCREMENT BY HOW MANY LINES;')
      IENDL=INUMB('END AT WHICH LINE;')
      ISTP=INUMB('START AT WHICH PIXEL;')
      INCP=INUMB('INCREMENT BY HOW MANY PIXELS;')
      IENDP=INUMB('END AT WHICH PIXEL;')
      CALL ASSIGN(7,'PUT RESULTS INTO WHICH FILE;',FILE)
      DO 200 I=1,NUMBND
  200 IVAL4(I)=0
      LNMOLD=-1
      II=512
       WRITE(7,10)
   10 FORMAT(' LINE PIXEL BAND WIDTH VALUES')
      DO 300 I=ISTL,IENDL,INCL
      DO 300 J=ISTP,IENDP,INCP
      II=512+322*(I-1)+J
      LNUM=(II-1)/8192
C      LNUM=(II-1)/512
      II=II-LNUM*8192
C     II=II-LNUM*512
      LNUM=1000*(2+LNUM)
      IF(LNUM.EQ.LNMOLD) GOTO 120
      LUN=7
      DO 310 K=1,NUMBND
      LUN=LUN+1
      CALL READ(LINE(I,K),LEN,2,LNUM,LUN)
  310 CONTINUE
      LNMOLD=LNUM
  120 CONTINUE
      DO 320 K=1,NUMBND
      CALL MOVEC(1,LINE(II,K),IVAL1(4,K))
  320 CONTINUE
      WRITE(7,20)I,J,(IVAL4(K),K=1,NUMBND)
   20 FORMAT(1X,1514)
  300 CONTINUE
      STOP
      END
      FUNCTION INUMB(MESS)
C **
C **  ENTRY TO RETURN NUMBER TO MESSAGE
C **
      LOGICAL*1 MESS(1),RESP(80)
      LOGICAL*4 QVAL
```

```
      INTEGER*2 LEN
C  ..  FIRST FIND SEMI COLONS
      ICNT=0
      CALL FINDST(MESS,100,';',1,1,IENDM,&124)
      IENDM=IENDM-1
      GOTO 150
  124 IENDM=10
  150 QVAL=.FALSE.
      ISTV=IENDM+1
C  ..  CHECK IF VALUES GIVEN
      CALL FINDST(MESS,100,';#',2,ISTV,IENDV,&200)
      IF(ISTV.NE.IENDV) GOTO 200
      CALL SETC(2,MESS(ISTV),' ')
      ISTV=IENDM+3
      CALL FINDST(MESS,200,';',1,ISTV,IENDV,&200)
      LENV=IENDV-ISTV
      IENDV=IENDV-1
      QVAL=.TRUE.
  200 ICNT=ICNT+1
      IF(ICNT.GT.5) GOTO 920
      IF(IENDM.GT.0)WRITE(6,10)(MESS(J),J=1,IENDM)
   10 FORMAT(1H&,100A1)
      CALL READ(RESP,LEN,0,LNUM,5,&900)
      IF(LEN.EQ.0) GOTO 200
C  ..  GET RID OF LEADING BLANKS
      ISTR=0
  340 ISTR=ISTR+1
      IF(ISTR.GT.LEN) GOTO 200
      IF(LCOMC(1,RESP(ISTR),' ').EQ.0) GOTO 340
C  ..  GET RID OF TRAILING BLANKS
      IENDR=LEN+1
  350 IENDR=IENDR-1
      IF(LCOMC(1,RESP(IENDR),' ').EQ.0) GOTO 350
      LENGTH=IENDR-ISTR+1
      LENG=LENGTH
C  ..  GET INTEGER VALUE
      CALL DTB(RESP(ISTR),INUMB,LENGTH,ISIG,' ',&200)
      IF(LENGTH.EQ.LENG) GOTO 360
C  ..  CHECK IF ILLEGAL CHARACTER IS A PERIOD
      ICH=ISTR+LENGTH
      IF(LCOMC(1,RESP(ICH),'.').NE.0) GOTO 354
      WRITE(6,44)
   44 FORMAT(' INTEGER VALUES ONLY ACCEPTED.  TRY AGAIN.')
      GOTO 200
  354 WRITE(6,48)
   48 FORMAT(' ILLEGAL CHARACTERS.  TRY AGAIN.')
      GOTO 200
  360 CONTINUE
      IF(.NOT.QVAL) GOTO 800
      ISTST=ISTV
C  ..  TEST FOR INTEGER
  240 CONTINUE
      IFELEM=0
  210 CONTINUE
      IFELEM=IFELEM+1
      IVAL1=IVAL(MESS,ISTST,LENV,'.-,:; ',6,ICH,IPOSN)
      ISTST=IPOSN+1
C  ..  IND=1(.)  IND=2(-)  IND=3(,)  IND=4(:)  IND=5(;)
      GOTO (900,220,230,900,230),ICH
  220 IVAL2=IVAL(MESS,ISTST,LENV,'.-,:; ',6,ICH,IPOSN)
```

```
121           ISTST=IPOSN+1
122           IF(INUMB.GE.IVAL1.AND.INUMB.LE.IVAL2) GOTO 800
123           IF(ICH.EQ.5) GOTO 900
124           GOTO 210
125         230 CONTINUE
126           IF(INUMB.EQ.IVAL1) GOTO 800
127           IF(ICH.EQ.5) GOTO 900
128           GOTO 210
129         800 RETURN
130         900 WRITE(6,40)(MESS(J),J=ISTV,IENDV)
131          40 FORMAT(' VALUE RECEIVED IS OUT OF RANGE.',
132            *  /,'RANGE MUST BE: ',80A1)
133           GOTO 200
134         920 STOP
135           END
136           SUBROUTINE ASSIGN(IO,MESS,FILE)
137       C   **
138       C   **  ROUTINES TO ASSIGN FILES AND SET BUFFER LENGTH
139       C   **
140           IMPLICIT INTEGER*4 (F)
141           LOGICAL*1 MESS(64),LINE(100),FILE(1)
142           LOGICAL*1 QNBUFF
143           INTEGER*4 ADROF
144           INTEGER*4 AREA(12),FDUB
145           INTEGER*2 LEN
146           EXTERNAL GETFD
147           INTEGER*4 RTN(6),FLAG,CINFO(16),FINFO(18),SINFO(6)
148           DATA RTN/6*0/
149           INTEGER*4 ERCODE
150           LOGICAL*4 ERRMSG(20)
151           COMMON/FI/FIAL,FIRL,FIFO,FIFLG,FICNS,FITS
152           COMMON/FI/FICPS,FIFLN,FILLN,FIMLL,FINE
153           COMMON/FI/FINL,FINH,FILCNT,FIHCNT,FIMHL
154           FIAL=16
155           QNBUFF=.FALSE.
156           CALL FINDST(MESS,60,';',1,1,IFND,&900,&900)
157           IFND=IFND-1
158           ICOUNT=0
159         100 ICOUNT=ICOUNT+1
160           WRITE(6,12)(MESS(J),J=1,IFND)
161          12 FORMAT(1H&,100A1)
162           ERCODE=0
163           IF(ICOUNT.GT.5) GOTO 900
164           CALL READ(FILE,LEN,0,LNUM,5)
165           IF(LEN.EQ.0) GOTO 100
166           LENP1=LEN+1
167           CALL SETC(1,FILE(LENP1),' ')
168           CALL GFINFO(FILE,RTN,1,0,0,0,ERCODE,ERRMSG,&103,&102)
169         102 CONTINUE
170           IF(ERCODE.EQ.0) GOTO 120
171           IF(ERCODE.EQ.2) GOTO 120
172       C   **  MUST CREATE FILE
173         103 WRITE(6,16)
174          16 FORMAT('FILE DOES NOT EXIST.')
175           ICODE=IOPROM('CREATE? NO OR YES:;','NO-YES;')
176           GOTO (100,104),ICODE
177         104 CALL FTNCMD('$CREATE ?',9,FILE)
178         120 CALL FTNCMD('ASSIGN ?=?',10,IO,FILE)
179           RETURN
180         900 WRITE(6,20)
```

```
181          20 FORMAT(1X,'IMPROPER ASSIGNMENT OF FILE')
182             STOP
183             END
184             FUNCTION IVAL(LINE,ISTIN,LENGTH,INKEY,LKEY,ICH,IPOSN)
185       C  **
186       C  **  FUNCTION TO GET VALUE DIRECTLY
187       C  **  RETURNS CODE INDICATING TYPE OF VARIABLE CAUSING STOP
188       C  **
189       C  ..  MAX LENGTH FOR NUMBER IS 10 DIGITS
190             LOGICAL*1 LINE(1),INKEY(1),KEYS(40)
191             KEYLEN=LKEY
192             IST=ISTIN
193             CALL MOVEC(KEYLEN,INKEY,KEYS)
194             CALL FINDST(INKEY,KEYLEN,'-',1,1,KMINUS,&120)
195             GOTO 130
196          120 KEYLEN=KEYLEN+1
197             KMINUS=KEYLEN
198             CALL MOVEC(1,'~',KEYS(KEYLEN))
199          130 CONTINUE
200             IF(LENGTH.GT.0) GOTO 90
201             CALL FINDST(LINE,20,':',1,1,IFIND,&80)
202             LENGTH=IFIND-1
203             GOTO 90
204           80 LENGTH=20
205           90 CONTINUE
206       C  ..  FIND DELIMITER KEY
207             ISTX=IST
208           94 ICH=0
209             CALL FINDC(LINE,LENGTH,KEYS,KEYLEN,ISTX,IPOSN,ICH,&99)
210             IF(ICH.NE.KMINUS) GOTO 100
211             IF(IPOSN.EQ.ISTX) GOTO 98
212             ILEN=IPOSN-ISTX
213             IF(LCOMC(ILEN,LINE(IST),'                    '))100,98,100
214           98 ISTX=IPOSN+1
215             GOTO 94
216           99 IPOSN=LENGTH+1
217          100 CONTINUE
218             IVAL=0
219             LEN=IPOSN-IST
220             LEN1=LEN
221             CALL DTB(LINE(IST),INT,LEN,IDNUM,' ',&999)
222             IF(IDNUM.EQ.0) GOTO 999
223             IVAL=INT
224             IF(LEN1.NE.LEN) GOTO 999
225             RETURN
226       C  ..  INDICATE RETURN ERROR
227          999 ICH=-1
228             RETURN
229             END
230             FUNCTION IOPROM(MESS,CHOICE)
231       C  **
232       C  **  RETURNS CODE OF RESPONSE
233       C  **  *MESS* AND *CHOICE* MUST END WITH A *;*
234       C  **
235       C  --  MESS IN FORM OF 'MESSAGE;'
236       C  --  CHOICE IN FORM OF 'CH1-CH2-CH3-CH4;'
237       C  --  IF NONE SELECTED THEN FIRST DEFAULTED TO
238       C  **
239             LOGICAL*1 MESS(100),CHOICE(100),RESP(100)
240             INTEGER*2 LEN
```

```
241           ICNT=0
242       C   **  FIRST FIND SEMI COLONS
243           CALL FINDST(MESS,100,';',1,1,IENDM,&121)
244           IENDM=IENDM-1
245           GOTO 122
246         121 IENDM=10
247         122 CALL FINDST(CHOICE,100,';',1,1,IENDC,&910)
248       C   ..  UPPER CHOICE
249       C     CALL UPPER(CHOICE,IENDC)
250          90 ICNT=ICNT+1
251           IF(ICNT.GT.5) GOTO 920
252           WRITE(6,10)(MESS(J),J=1,IENDM)
253          10 FORMAT(1H8,100A1)
254           CALL READ(RESP,LEN,0,LNUM,5,&900)
255           LENGTH=LEN
256       C   ..  UPPER RESPONSE
257       C     CALL UPPER(RESP,LENGTH)
258           IF(LEN.EQ.0) GOTO 90
259           ICODE=0
260           ISTOPC=0
261           IFND=0
262         100 CONTINUE
263           IF(IFND.GE.IENDC) GOTO 90
264           ISTC=ISTOPC+1
265           ICODE=ICODE+1
266           CALL FINDC(CHOICE,IENDC,'-;',2,ISTC,ISTOPC,IFND,&90,&90)
267           CALL FINDST(CHOICE,ISTOPC,RESP,LENGTH,ISTC,IFNDC,&100)
268           IF(IFNDC.NE.ISTC) GOTO 100
269           IOPROM=ICODE
270           RETURN
271         900 IOPROM=1
272           RETURN
273         910 WRITE(6,20)
274          20 FORMAT(1X,'ERROR IN MESSAGE OR CHOICES')
275           RETURN
276         920 STOP
277           END
```

APPENDIX 4

Control points and residual values from the affine transformation on the geometrically-corrected image

UTM Coordinates

Image Coordinate		Map Easting (m)			Map Northing (m)		
Pixel	Line	Actual	Estimated	Residual	Actual	Estimated	Residual
35	401	445325.00	445419.94	-94.94	6652987.00	6653012.50	-25.50
97	328	447238.00	447259.91	-21.91	6655225.00	6655192.50	32.50
127	429	448237.00	448164.12	72.88	6652212.00	6652166.50	45.50
275	440	452650.00	452575.28	74.72	6651763.00	6651825.50	-62.50
309	257	453500.00	453569.69	-69.69	6657275.00	6657301.00	-26.00
214	317	450725.00	450745.06	-20.06	6655600.00	6655512.50	87.50
7	188	444575.00	444563.81	11.19	6659300.00	6659391.00	-91.00
28	70	445225.00	445177.53	47.47	6662975.00	6662922.00	53.00
169	26	449375.00	449374.47	0.53	6664213.00	6664227.50	-14.50

Standard error of the pixel estimate 68.338m

Standard error of the line estimate 67.448m

APPENDIX 5

SPSS version 9 command files

for geographic referencing

```
run *spss 6=-out
run name          AFFINE TRANSFORMATION - THESIS
variable list     x,y,x1,y1
input med         card
n of cases        9
input format      fixed(2x,f4.0,3x,f5.0,3x,f7.0,3x,f8.0)
var labels        x,PIXEL/y,LINE/
                  x1,UTM EASTING/y1, UTM NORTHING/
new regression    descriptives/
                  variables=x to y1/
                  statistics=defaults,ci,label,f/
                  dep=x1/enter x,y/
                  residuals/
                  casewise=all, dependent,pred,zresid/
                  dep=y1/enter x,y/
                  residuals/
                  casewise=all,dependent,pred,zresid/
read input data
   35.    401.    445325.    6652987.
   97.    328.    447238.    6655225.
  127.    429.    448237.    6652212.
  275.    440.    452650.    6651763.
  309.    257.    453500.    6657275.
  214.    317.    450725.    6655600.
    7.    188.    444575.    6659300.
   28.     70.    445225.    6662975.
  169.     26.    449375.    6664213.
finish
```

Program to compute affine transformation equations

```
1     run *spss 6=-out
2     run name          AFFINE TRANSFORMATION - THESIS & UTM
3     variable list     L,P
4     input medium      disk(band)
5     n of cases        198
6     input format      fixed(1x,2f4.0)
7     compute           X=444336.06426 + 29.79737*P + .10218*L
8     compute           Y=6665019.44 - .08018*P - 29.93643*L
9     var labels        P,PIXEL/L,LINE/
10                      X, EASTING/Y, NORTHING/
11    list cases        forced/cases=198/variables=L,P,Y,X/
12    finish
```

Program to compute UTM coordinates -